DOCUMENT RESUME

ED 146 177 TM 006 509

AUTHOR Veroff, Joseph; And Others
TITLE Measuring Intelligence and Achievement Motivation in Surveys. Final Report to U.S. Department of Health, Education and Welfare. Office of Economic Opportunity.
INSTITUTION Michigan Univ., Ann Arbor. Survey Research Center.
SPONS AGENCY Office of Economic Opportunity, Washington, D.C.
PUB DATE Oct 71
CONTRACT OEO-4180
NOTE 132p.; For related documents, see TM 006 510 and 511

EDRS PRICE MF-$0.83 HC-$7.35 Plus Postage.
DESCRIPTORS Achievement; Age Differences; Educational Background; Evaluation Criteria; Field Interviews; Intelligence; *Intelligence Tests; Mediation Theory; *Motivation; Predictor Variables; *Psychological Tests; Questionnaires; Race Influences; Racial Differences; Research Methodology; Risk; Sex Differences; Social Differences; *Surveys; Test Bias; *Test Construction; Test Reliability; Test Validity

ABSTRACT

A series of studies and analyses to develop brief, reliable measures of intelligence and achievement motivation that would be feasible for use in a cross-sectional sample is summarized. One measure of intelligence meets these criteria. It is a 13-item test of sentence completion skill, which correlates effectively with other measures of intelligence and with status and achievement variables in a meaningful way. A questionnaire measure of achievement motivation also emerges as a potentially useful measure for most groups except black females. It correlates moderately well with both a projective measure of achievement motivation and a behavioral assessment of moderate risk-taking in most groups, except the black female group. Both the intelligence and the motivation measures suggested are independently effective in predicting the education attained by black males and white females regardless of their status backgrounds. Only the intelligence measure is an independent predictor of education attainment in white males and black females, regardless of their status backgrounds. (Author/MV)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED146177

~~LN-1386~~

TM

Final Report to

U.S. Department of Health, Education and Welfare
Office of Economic Opportunity
Contract No. OEO-4180

MEASURING INTELLIGENCE AND ACHIEVEMENT MOTIVATION IN SURVEYS

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Joseph Veroff, Lou McClelland,
and Kent Marquis

Survey Research Center
Institute for Social Research
University of Michigan
Ann Arbor, Michigan

October 1971

The research reported herein was performed pursuant to a contract with the Office of Economic Opportunity, Executive Office of the President, Washington, D.C. 20506. The opinions expressed herein are those of the authors and should not be construed as representing the opinions or policy of any agency of the United States Government.

TM006509

TABLE OF CONTENTS

Page

SECTION I

ABSTRACT . . . .

ACKNOWLEDGEMENTS . . . .

LIST OF TABLES . . . .

LIST OF FIGURES. . . . .

I. INTRODUCTION 1

The Problem 1
Plan of this Monograph 4

II. BACKGROUND TO THE DETROIT STUDY 6

The Jackson I Study 6
The Jackson II Study 8

Measuring Motivation 8
Relationship of Questionnaire Measures of Motivation to Behavioral Measures 13
Measuring Intelligence 14

Where We Stood Prior to the Detroit Survey 22
The Detroit Survey: Overall Procedure 23
The Sample 24

Measures 24

III. VALIDATING MEASURES OF INTELLIGENCE 26

Measures 26

Analysis 28

Intercorrelations of Measures of Intelligence 29
Correlations of Intelligence Measures with Education . . 38
Correlating the Sentence Completion Test with other Tests, Partialing Out Demographic Correlates 40
A Measure of Intelligence to be Used in Surveys: the Sentence Completion Test 41

Use of the Test 45

Page

IV. METHODS OF ASSESSING ACHIEVEMENT MOTIVATION IN A SURVEY: THEORETICAL BACKGROUNDS AND DESIGN OF STUDY . . 48

Plan of the Interview 53
Practical Considerations 56
Plan of Analysis 58

V. METHODS OF ASSESSING ACHIEVEMENT MOTIVATION IN A SURVEY: ANALYSIS OF DATA 59

Criteria Development 59

Projective Criterion 59
Risk-Taking Criterion 63

Component Development 71

A Priori Component Measures 71
Empirically-Derived Component Measures 71

Testing the Scales 83

A Priori Scales 83
Empirical Scales 83

White males 89
Black males 89
White females 92
Black females 94

Empirical and A Priori Scales Compared: A Summary 94
Increasing the Effectiveness of the A Priori Scales 96

A new test anxiety scale 96
New scale for achievement orientation 98

A Resultant Achievement Orientation Scale 98
Summary 104

BIBLIOGRAPHY 108

SECTION II

APPENDIX 1: THE INTERVIEW

APPENDIX 2: MEASURES

SECTION III

APPENDIX 3: FACTOR ANALYSIS OF MEASURES OF ACHIEVEMENT MOTIVATION

APPENDIX 4: RESULTS OF THE INTERVIEWER-RESPONDENT RACE EXPERIMENT

APPENDIX 5: TWO METHODOLOGICAL STUDIES

APPENDIX 6: SOME PROVOCATIVE RESULTS WITH THE PICTURE ORDER INCIDENTAL TEST

ABSTRACT

This report summarizes a series of studies and analyses carried out to develop brief measures of assessing intelligence and achievement motivation that would be feasible in a cross-sectional sample and reliably reflect what has been standardly assessed in these domains by past researchers. One measure of intelligence stood out above others in meeting our criteria. It is a 13-item test of sentence completion skill. It correlates effectively with other measures of intelligence and with status and achievement variables in a meaningful way. A questionnaire measure of achievement motivation emerges as a potentially useful measure in most groups except black females, one that correlates moderately well with both a projective measure of achievement motivation and a behavioral assessment of moderate risk-taking in most groups, except the black female group. Both the intelligence and the motivation measures suggested are independently effective in predicting the education attained by black males and white females regardless of their status backgrounds. Only the intelligence measure is an independent predictor of education attainment in white males and black females, regardless of their status backgrounds.

ACKNOWLEDGEMENTS

The impetus for the research reported in this monograph came from James Morgan, James Smith, and the late John Lansing, who were interested in the results of our work for their larger project aimed at understanding the dynamics of changes in poverty status. They helped to structure the problem clearly and yet generally supported our desires to extend the research beyond their pragmatic goals.

A number of people served as active researchers at different stages of our work. John Atkinson, Charles Cannell, Grace Mack, and David Ruhland participated in our initial studies, what might be viewed as pretests to the major study described in our monograph. Andrew Bustin, Ateleba Crespi, and Michael Stinson served as research assistants during various stages of the major study. Most important, however, was the assistance of Robert Hubbard at all stages of this work. His labors and wise counsel are deeply appreciated.

John Hagen and Gerald Gurin offered valuable suggestions at one time or another. A number of people helped in data analysis: Margie King, Sherry Lucas, Diane Schoeff, and Linda Winter. We would also like to acknowledge the pleasure of their company. Our typists, Roberta Petersen and Marion Wirick are extraordinarily competent at reading hieroglyphics.

LIST OF TABLES

Number	Title	Page
2.1	Different Types of Questionnaire Items for Measuring Components of Achievement Motivation, with Examples . . .	9-10
2.2	Behavioral Measures of Motivation used in the Jackson II Study	12
2.3	Correlations (Tau Beta) of Digit Symbol Substitution Test with Cloze and Maze Tests by Respondent Demographic Group	17
2.4	Correlations (Tau Beta) of Sentence Completion Test (Lorge Thorndike) with Cloze and Maze Tests by Respondent Demographic Group	18
2.5	Correlation (Tau Beta) Between Digit Symbol Substitution Test and All Other Tests	20
2.6	Correlation (Tau Beta) Between Sentence Completion Test and All Other Tests	21
3.1	Intelligence Measures Used in the Detroit Survey	28
3.2	Mean Scores on Intelligence Measures, for Eight Population Groups	31
3.3	F. Ratios from the Analyses of Variance on Eight Intelligence Measures: Race of Interviewer x Race x Sex of Respondent	32
3.4	Mean Education Attained, by Interviewer and Respondent Race	33
3.5	Intercorrelations of Intelligence Measures, by Respondent Race	35
3.6	Intercorrelations of Intelligence Measures, by Socio-Economic Background Groups	37
3.7	Correlations of Education with Intelligence Measures . . .	38
3.8	Correlations of Education with Intelligence Measures, by Race and Socio-Economic Background Groups	39
3.9	Intercorrelations of Intelligence Measures, with Education, Age, Race, Sex, Interviewer Race, and Region Partialed Out	41

Number	Title	Page
3.10	Sentence Completion Test	42-44
5.1	Intercorrelations of Projective Assessments of Achievement Motivation, by Sex	60
5.2	Analysis of Covariance of Achievement Imagery on P1, P2, and P4, with Length of Response as Covariate; Race of Respondent x Race of Interviewer x Sex	61
5.3	Regression on Achievement Imagery on P1+P2+P4, Race of Interviewer and Length of Response as Predictors	62
5.4	Intercorrelations of Behavioral Risk-Taking Measures, by Sex	64
5.5	Percent of Respondents Choosing All "Very Difficult" Choices, by Sex and Intelligence Measures	66
5.6	Correlation between I.Q. (Sentence Completion Test) and Mean Difficulty of Choice, by Sex and Race Groups	67
5.7	Analysis of Covariance of Mean Difficulty in Risk Preference, with I.Q. as Covariate Race of Respondent x Race of Interviewer x Sex	68
5.8	Regression on Mean Difficulty of Choice, Interviewer Race and I.Q. as Predictors	69
5.9	Relation of Risk-Taking Criterion, Corrected Risk Difficulty, and Uncorrected Risk Choice Scales	70
5.10	A Priori Achievement Motivation Scales	72
5.11	Past Test Anxiety Scale	74
5.12	Personal Efficacy Scale	75
5.13	Coefficients of Internal Consistency (α) for the A Priori Scales	76
5.14	Items Used in Factor Analysis, with Abbreviations	77
5.15	Empirical Scales Resulting from Factor Analyses, White Males	84

Number	Title	Page
5.16	Empirical Scales Resulting from Factor Analyses, Black Males	85
5.17	Empirical Scales Resulting from Factor Analyses, White Females	86
5.18	Empirical Scales Resulting from Factor Analyses, Black Females	87
5.19	Correlations of Empirical and A Priori Scales with Risk-Taking and Projective Criteria, by Race-Sex Groups	88
5.20	Items with Loadings of .40 or Over on Empirical Scales Effective for White Males	90
5.21	Items with Loadings of .40 or Over on Empirical Scales Effective for Black Males	91
5.22	Items with Loadings of .40 or Over on Empirical Scales Effective for White Females	93
5.23	Items with Loadings of .40 or Over on Empirical Scales Effective for Black Females	95
5.24	Revised Anxiety Scale with Alpha Coefficients and Correlations to Criteria in Four Race-Sex Groups	97
5.25	Correlations of Item-to-Criteria for Items in Three A Priori Scales, by Race-Sex Group	99
5.26	Achievement Orientation Scale	100
5.27	Achievement Orientation Scale: Means, Sigmas, Alphas and Correlations to Criteria in Four Race-Sex Groups	100
5.28	Resultant Achievement Orientation Scale: Alpha Coefficients and Correlations to Criteria in Four Race-Sex Groups	101
5.29	Correlations of Resultant Achievement Orientation to Selected Characteristics of Respondents (by Sex-Race Group)	101
5.30	Correlation of Education to Projective and Risk-Taking Criteria (by Sex-Race Groups)	103

Number	Title	Page
5.31	Correlations of Resultant Achievement Orientation (A) to the Projective and Risk-Taking Criteria (C), Partialling out the Effects of Education (E) in Different Sex-Race Groupings	103
5.32	Regression on Education with Intelligence (SCT) and Resultant Achievement Orientation as Predictors along with Status Background and Age (By Sex x Race Groups)	105

LIST OF FIGURES

Number	Title	Page
3.1	Distribution of the Scores Obtained on the Sentence Completion Test	46

CHAPTER ONE

INTRODUCTION

In early 1971, 365 Detroiters took part in a household survey that attempted to assess intellectual competence and achievement motivation. The major portion of this monograph represents a review of that survey -- how it came into being, its procedures, major results, and implications. From our way of thinking, this study is more than a significant contribution to understanding appropriate methods of assessing intelligence and motivation in a survey setting; it also contains some special advances in our understanding of the nature of intelligence and motivation. Although the research was contracted only to answer pragmatic assessment questions, we think we have some new conceptual insights about American social phenomena. Therefore, this report will point to the methodological, conceptual, and social conclusions that stem from the Detroit survey.

The Problem

The impetus for these investigations came from Morgan and Smith's long term and continuing interest in using psychological factors as explanatory variables in considering economic behavior (Morgan and Smith, 1970). Variables such as the ones we are exploring in this report -- intelligence and motivation -- can be viewed in two ways. They can lend new psychological information about sources of economic behaviors since they are relatively independent of the socio-economic variables usually employed in macro-sociological analyses. Or they can be thought of as intervening variables linking the macro-sociological variables with economic behavioral variables. Such intervening variables lend a richer picture to the explanatory network but do not have an independent status as predictors. In either case, insights about these psychological variables can aid in conceptualizing what effects various programs of planned economic change might have. In this report we are assuming that measures of intelligence and motivation are exciting variables to include in survey studies and that our task is to come up with the best way to measure them.

The methodological work we are about to present had the principle aim of developing valid measures of motivation and intelligence which could be used in household interviews and could meet the following criteria:

1. Be feasible in a cross-section sample of the United States population;
2. Be reliable and valid for major groups within the population;
3. Not provoke hostility or anxiety;
4. Be extremely brief (not more than five minutes even for respondents difficult to interview).

These are fairly stringent criteria for survey methods applied to such complicated psychological variables as motivation and intelligence; perhaps the most severe is the criterion of validity. As soon as a social scientist asks whether a measure of a complex psychological characteristic is valid not only for a special group of the population but for all major groups of a population, he is confronted with one of the most challenging tasks he can set for himself. In so doing he can foresee two important considerations: (1) there is some likelihood that no single measure of intelligence or motivation could do the explanatory job; (2) there is no single criterion for validity of a given measure. These two considerations are not unrelated. As soon as a scientist owns up to a multiplicity of measures for a conceptual area of investigation, in all likelihood he is faced with a set of criteria for evaluation of validity rather than just a single one. Let us discuss the problem of multiple measurement first.

From Atkinson's theory and research on achievement motivation (1964), we knew that it would be foolhardy to expect a single measure of achievement motivation* alone to have much predictive value. Atkinson's work uses a motive for success in interaction with a motive to avoid failure, and also considers expectancies and incentives quite separately from motives. Other achievement motive theorists have pointed to the importance of still other components of the general motive; these are discussed fully in Chapter 4. Hence we saw that our task might be to come up with not a single measure of motivation, but multiple measures. Given the severe time limitation for assessment of motivational variables, our task became even more demanding.

---

*We will follow Atkinson's terminology, using "motivation" to refer to the joint effect of various determinants of behavior and "motive" and "expectancy" to refer to specific components of over-all motivation. At times we will use the term "orientation" as a generic term for various types of components of motivation.

Recent concerns about cultural bias in measurement of intelligence had underscored the possibility that there should also be a multiple criterion approach to the assessment of intelligence. Particularly striking to us was the possibility of investigating two intellectual factors: one that specifically reflects the effects of learning the dominant cultural intellectual demand -- a verbal mediational facility; and another reflecting a less obviously culturally biased facility -- a perceptual performance facility. This is a distinction traditionally made, and we thought it worth keeping for the concerns of OEO. So again, we took it as a challenge to develop at least two different assessment procedures for intelligence, wishing that there were time in the panel study for even more. As we shall see, we ultimately found a single measure sufficient for current purposes. Nevertheless, the research design anticipated multiple measurement.

How can the validity of each of several multiple measures be assessed? We approached this problem in three ways. First, we took existing measures as anchors, with new measures evaluated on their correlations with these anchors. Secondly, we selected a particular behavior and asked if a new measure correlated with that behavior. Finally, we used the one approach to validity that today has the most currency in social science -- construct validation. In this approach a measure must be related to a whole network of other variables that make some sense in a theoretical scheme.

It was quickly apparent in approaching the methodological investigation of multiple measures of intelligence and motivation that both fields of research have had divergent approaches to validity of measurement. The measurement of intelligence has relied on very pragmatic criteria of validity -- particularly educational success. The more theoretical approaches to intelligence (e.g., Piaget) have included little systematic measurement of individual differences. Thus very refined measurements have been used for intelligence without the kind of theoretical underpinning that could guide a scientist who wished to adopt the strategy of construct validation in methodological investigation. We continued the usual tradition of validating intelligence measures in relationship to standard tests and to educational attainment.

By sharp contrast, there has been considerable theoretical development in the theory of achievement motivation, but less work on techniques for assessment of individual differences. We thus had guidelines for studying construct validity but little to turn to in previously refined assessments of motivation to anchor our work. We therefore used both a theoretically derived measure of risk-taking and an empirically derived projective fantasy measure as criteria for establishing the validity of motivational measures.

With multiple measures and complex criteria for analyzing validity, we proceeded cautiously. There were no easily agreed upon guideposts for tracking down the validity of a given measure. The ins and outs of the progress we are about to present by necessity did not always follow clear paths. We found that, although we followed specific strategies for validation, in the course of implementing these strategies we came across new findings, new ways of looking at data that made us proceed on a somewhat ad hoc basis. As a result any recommendations for use of techniques in the future should be considered as cautious appraisals lacking overwhelming scientific underpinnings.

Plan of this Monograph

In planning the Detroit survey we had to go through a number of preparatory steps: first, review the available assessment procedures for intelligence and motivation in order to select techniques feasible in a survey setting; second, design strategies to establish the validity of these measures; and finally, pretest the measures to assure their feasibility. The second chapter deals with these background issues, and presents the design of the Detroit survey. Following are chapters summarizing the results and recommendations for the intelligence and motivation measures. Appendices are attached covering some specific studies and methodological information which tend to interrupt the flow of presentation to the general reader, but may be of interest to researchers actively working on these problems. Appendices including the interviewers' manual, the respondent's booklet, and the actual test materials used in the interview are available from the authors on request. Limited interest and copyright laws prevent their general distribution.

CHAPTER TWO
BACKGROUND TO THE DETROIT STUDY

Measures of intelligence and achievement motivation which could be used in household interviews with a cross section of the population were needed for incorporation into a five-year panel study of variables influencing movement in and out of poverty. Several pilot studies investigating the usefulness of existing intelligence and motivation measures were undertaken and are described below. The results of these studies pointed up the complexity of the measurement problem and highlighted areas in need of more intensive study with larger numbers of respondents. The major results of the pilot studies which shaped the design of the final research are given here.

The Jackson I Study

In directing the first years of the panel study of changes in economic status, James Morgan was interested in having available quick objective measures of achievement motivation that would be useful for predicting economic change. The Survey Research Center has been engaged in methodological investigation of measures of achievement motivation in a national survey for many years (Veroff _et_ _al_., 1960; Morgan and Smith, 1970). Recently Atkinson and Moulton (1969) started on the first step of a concerted effort to come up with objective measures of motivation that could be generally useful. Consequently, Morgan enlisted Atkinson and his colleagues to conduct a household survey to see whether the methods emerging from their work with college students could be easily transferable to a more heterogeneous household sample. During the year 1968-69 Atkinson directed such a study in Jackson, Michigan (Jackson I Study) that began our methodological work for OEO. Previously validated methods were adopted and new methods explored in studying 50 adults selected on a quota basis (half black, half white; half male, half female) from a relatively low socio-economic area of Jackson for the survey. While this study showed that the methods could be applied, and that there was no difficulty in getting respondents to give responses, there was little internal consistency within the set of items used, especially in the most promising scale of achievement motivation developed with college students by Atkinson and Moulton.

Some more specific results of this study proved more interesting and challenging than originally anticipated: (1) The order of presentation of open-ended questions affected the quality of the response (e.g., when a question about jobs was asked first, twice as much codable economic imagery was found as when the question was asked later in the interview); (2) respondents were generally unwilling and/or unable to give more than four sets of fantasy-like responses (e.g., telling a story to a verbal lead); (3) proper probing techniques on the open-ended questions were difficult to specify.

Some analyses of objective scales were encouraging for future research. Comparison of mean scores on Atkinson and Moulton's objective scale with mean scores for a Flint college night school class showed higher means and a more restricted range for the Flint sample. The pattern of intercorrelations was also different, with a Personal Efficacy Scale correlating higher with the objective motivation scale in the Flint sample than in the Jackson sample.

All of the methodological and content results, taken together, indicated that: (1) This type of questionnaire approach was feasible and justified further research, although some individual items might not be clearly understood; (2) behavioral validity criteria would help in the evaluation of the questionnaire measures; and (3) as we moved to a more heterogeneous population the multidimensionality of achievement motivation became more important to consider. At the conclusion of the Jackson I Study, Joseph Veroff was asked to take over Atkinson's position, since Atkinson was out of the country for a year.

The Jackson II Study

Guided by the results of the Jackson I Study, we were eager to try out different and less abstract ways of wording items. We also began thinking seriously of different conceptual facets of achievement motivation. Because Jackson I lacked any specific behavioral bases for validating the methods and dimensions examined within the survey itself, we wanted to include this type of measure. These became the major aims in our methodological study of motivation in the second Jackson Study (Jackson II).

In Jackson II we also began systematic methodological exploration of measures of intelligence. We will discuss motivation and intelligence separately.

Measuring Motivation

With the Jackson II Study we systematically began to isolate components of achievement motivation. From other researchers' examination of different groups (Epps, 1969 on black high schoolers; Gurin, et al., 1969 on black college students; Sommerfeld, 1969 on job trainees; and Atkinson and Moulton, 1969, on college students) we considered items that fit slightly different conceptual dimensions of achievement motivation: future orientation, avoidance of social approval, social comparison for achievement, strong emphasis on autonomy, preference for moderate risk.

Table 2.1

Different Types of Questionnaire Items for Measuring Components of Achievement Motivation, with Examples

Component of Achievement Motivation	Example
Risk Orientation	Imagine doing something where your chances of success are 50-50. a. Would you rather be doing something else? b. Do you enjoy the challenge and competition?
Future Orientation	When you finish something important to your future plans, what usually happens next? a. You start working on other things for the future. b. You relax and are glad not to have to worry about doing the work again.
Interest in Social Comparison	When you try to get better at something, do you a. See how you do compared to other people, or b. See if you have improved since the last time you tried to do it?
Self-blame for Failure	When you can't do something that you really want to do, which is most likely to happen? a. You think about why you failed and try again. b. You forget about failing and don't try again.
Concern about Social Approval for Achievement	After you do a job well, which comes closest to describing why you feel so good? a. Because you will get praise for working hard. b. Because you know you did something well.
Self-esteem about Achievement	"I feel I do not have much to be proud of." a. Agree b. Disagree

Table 2.1 -- Continued

Component of Achievement Motivation	Example
Salience of Achievement for Self	Which is truest for you? Which of the three is least true? a. I would like to have more friends. b. I would like to do better at what I try. c. I would like to have more people pay attention to my point of view.
Personal Efficacy (1st Person)	"I am able to do things as well as most other people." a. Agree b. Disagree
Personal Efficacy (3rd Person)	"It doesn't make much difference what a person tries to do; some folks are just lucky, others are not." a. Agree b. Disagree
Competence Orientation to Feelings of Pride	There are times when many of us are proud of something we have done. Tell me about some time when you have been proud of something you have done. (With further probes about a specific example)
Admission of Failure	Tell me about a time when you have felt dissatisfied with something you have done. (Similar to above)
Achievement Orientation to Failure	(In probes for above question) Why did this make you feel bad?
Retrospective Report about Anxiety in Tests	Test Anxiety Questionnaire: "My emotions hurt my performance on tests I took in school." a. Agree b. Disagree

These are further described in Appendix 2. Although we were interested in how these measures related one to another, we were especially interested in how any of the above related to the behavioral measures to be discussed below. From the Jackson I Study we realized we were lacking any behavioral criteria for estimating the validity of proposed measures of achievement motivation. We therefore set up what we think are innovative techniques for getting at standardized assessment of risk preference behaviors as well as performance on achievement tasks under different probabilities of success. These techniques, described in more detail below, tap risk preference and performance under various expectations of success, behaviors that have been criterion variables used by Atkinson and Moulton, and Veroff in other contexts. These criteria then became validity checks for evaluating questionnaire measures of achievement motivation.

A series of behavioral tasks derived from Atkinson and Feather's (1966) thinking were given to the respondents in the Jackson II survey by graduate students familiar with current methodology in achievement motivation research. The principal measures were built around performance on and reactions to some of the intelligence measures. We felt that we had a unique opportunity to get what we considered motivationally relevant reactions to the respondents' own performance on the intelligence tasks. We were particularly interested in the respondents' choices of moderately difficult performance (Risk Preference 1 and Risk Preference 2 are described in Table 2.2 below.) and the respondents' differential performance on hard, moderately hard, and easy tasks (Reactive Performance and Performance under Different Perceived Difficulties). Table 2.2 summarizes the behavioral measures of motivation used and the type of performance (usually intelligence) instrument in which this motivational assessment was set.

Table 2.2

Behavioral Measures of Motivation used in the Jackson II Study

Measure	Performance Setting	Index of High Achievement Motivation
Reactive Performance (Differential effects of success and failure on subsequent performance)	Performance on Digit Symbol Substitution following success and failure at Porteus Mazes (described in Appendix 2)	Performance after failure minus performance after success
Risk Preference 1: Repeated Observations (Repeated observations of choosing moderate risk in preference to easy or difficult choices)	Selecting line puzzles of different difficulty to work on (Weiner, 1963)	Number of moderate choices
Risk Preference 2: Single Observations (Choosing moderately hard task after failure at that task)	Digit Span and Block Design (described in Appendix 2)	Moderate
Performance under Different Perceived Difficulties (Performance at a moderately hard task contrasted to performance at an easy and a difficult task)	Three forms of the Digit Symbol Substitution Task varying in length to be performed in constant time	Performance at a moderately hard task - 1/2 (performance at easy task plus performance at hard task)

Because we wanted a reliable behavioral index, we looked for a composite of the separate indices of motivation from the behavioral criteria. However, the intercorrelations of the various measures for the total sample were low and there were very few positive trends relating pairs of these measures. When the sample was divided by sex, however, the results for males of both races did replicate some of those obtained with a college male sample. Sometimes the female results showed an opposite pattern. As other results were investigated, this differential male-female pattern continued. And so we thought we were on to something systematic. Due to the restricted sample size, however, adequate analyses of the results were not feasible. Hypotheses, especially those dealing with male-female differences, could only be tested in a larger sample where more detailed analysis would be possible.

Relationship of Questionnaire Measures of Motivation to Behavioral Measures

As in the Jackson I Study, there was little to indicate that various components of the questionnaire measure of motivation were consistently related to each other. There were again some widely divergent interrelationships when the data were analyzed separately for men and women. What became particularly clear is that responses to even mildly abstract references to achievement (e.g., "challenging task") were just not related to any criteria of validity that we used. The items that did work seemed to have more specific concrete references, such as to a person's job. Any further valid assessment in a heterogeneous population would therefore have to take that into account. We found that rank ordering long lists of items might not be meaningful, but that rank ordering three items did produce some meaningful correlations to behavioral criteria. Furthermore, some of the open-ended questions included were promising in relation to behavioral criteria, and we therefore reconsidered open-ended strategies for the future.

Dimensions of achievement motivation seemed to work differently in different groups: for example, social comparison seemed to be a positive achievement orientation in women but a negative achievement orientation in men. Orientation to the future seemed to be more diagnostic for women than for men. A high general expectancy of efficacy was more diagnostic for men than for women. Thus, scales that looked promising for one subgroup either had low correlations for the other subgroups or were opposite in sign. The preferred psychometric technique for identifying meaningful dimensions would have been factor analysis, but the subgroup size was not sufficient for this procedure. The larger sample size we planned for the Detroit Study partially came out of the idea of using factor analysis programs on subgroup data separately for men and women.

Measuring Intelligence

The second part of the Jackson II Study was the beginning of an evaluation of techniques to be used to measure intelligence in a nationwide survey. A short (5-minute) valid measure was wanted for the panel study. Previous work with the Ammons Quick Test in a small sample of the panel study was encouraging and an in-depth look at such a measure was wanted. We were given to understand that we should get feasible ways of measuring what psychologists have customarily called intelligence. Although it was tempting to use new criteria for intelligence, especially taken from real life behavior (i.e., successful performance at a job, income or occupational mobility) we were asked to avoid such criteria. The investigators in the OEO project wanted to know whether or not the intelligence measures were useful in predicting economic behavior. If methodological evaluation were itself based on real life economic criteria we were virtually guaranteeing that the measures would be correlated with changes in economic status. We took as our main charge, therefore, getting tests that correlate with standardized assessment procedures. This is certainly a traditional approach to validity, but one that has some justification in this case, especially because the theoretical underpinnings to the nature of intelligence are weak in systematic research investigation, as we previously noted.

Within that overall mandate, however, we felt we did have some leeway. Traditional intelligence tests have not always seemed to measure the same thing, for measurement theorists think of intelligence in a variety of ways. Some assume a g factor, a core intelligence related to other more specific competences. Others talk of components without any g factor. Our original orientation was to assume separable components -- particularly verbal mediational abilities contrasted to perceptual performance abilities, as mentioned earlier. Within each of these clusters we hoped to be able to select a single most feasible measure, especially one that had strong face validity.

Immediately a question arises: is the same measure equally valid for all groups? It certainly simplifies consideration to think so, but that belies some of our best social psychological insights. Subcultural patterns in different groups can channel general competences into very different kinds of specific performances. We were hoping to find tests that would be equally valid for the following groups -- men and women, blacks and whites, old and young. Nevertheless, we wanted to be alerted to the possibility that such tests may not be available and advise OEO accordingly.

In summary, we looked for validity in measures of intelligence by: hypothesizing two clusters of abilities, a verbal mediational facility and perceptual performance facility; then looking at the patterns of intercorrelations within the cluster and seeing whether the intercorrelations were consistent across critical groups, and finally seeing that the measures had face validity for a nationwide survey.

Between April 1969 and June 1969 a group of us surveyed available material on measuring intelligence. Two major American survey techniques include a multiple choice-vocabulary type of test (Ammons Quick Test and Miner's Vocabulary Test). These were considered along with subtests of the most widely used American tests -- the Weschler-Bellevue and the Stanford Binet. (See Appendix 2 for a complete description of the measures used.) We also looked at and finally adopted the Porteus Maze Test (purported to be relatively culture free and reflecting a series of skills, the most important of which is pre-planning); a Cloze procedure recommended by Jensen to assess what he thinks is the predominant factor in most intelligence measures of abstract thinking -- verbal mediation; a more standardized version of the Cloze procedure found in the Lorge-Thorndike test; and a multiple choice fill-in of incompleted sentences (henceforth called the Sentence Completion Test). After consulting with local testing experts, we selected the following tests from the Weschler-Bellevue and from the Stanford-Binet:

Similarities - (e.g., how are oranges and apples alike?)
Information - (e.g., what is a thermometer?)
Block Design - (arrange blocks to duplicate patterns)
Digit Span - (repeating sequences of digits backwards)
Digit Symbol Substitution - (writing down symbols encoding numbers according to a pre-arranged code)

Since the population in the Jackson II Study had a limited range of educational experience, we could not use educational attainment as a criterion for testing the validity of measures of intelligence, as we would in the Detroit Survey. Rather we had to adopt certain measures as a priori valid measures and then assess the degrees to which other measures correlate with these standards. Since we had hypothesized a two-factor assessment of intelligence, we selected two criteria.

A priori, the Maze test was selected as the best measure of performance and the Cloze test as the best verbal measure. With the adoption of these criteria, we could ask which of the remaining tests best predicts both criteria, or which test predicts the first criterion, and which the second. These questions were further complicated by an indication that a prediction to a criterion was not stable across all groups -- male and female, black and white, older and younger.

Asking which test reliably predicts both Cloze and Maze tests, we found that the Digit Symbol Substitution Test (DSST) was best, and was reliable across all groups except people over 35 (see Table 2.3 on page 17). Asking which test reliably predicts Cloze performance, we had a very clear answer -- the Sentence Completion Test (see Table 2.4 on page 18). Asking which test reliably predicts the Maze performance, we had a clear answer again -- the Digit Symbol Sutstitution Test (DSST).

Table 2.3

Correlations (Tau Beta) of Digit Symbol Substitution Test with Cloze and Maze Tests by Respondent Demographic Group

Respondent Demographic Group	N	Cloze	Maze
By Sex:			
Males	25	.43*	.32
Females	24	.33	.25
By Race:			
White	25	.40*	.37
Black	24	.34	.17
By Age:			
Over 35	25	.38*	.04
Under 35	24	.28	.52**
TOTAL	49	.36**	.25

*p = .05.

**p = .01.

Table 2.4

Correlations (Tau Beta) of Sentence Completion Test (Lorge Thorndike) with Cloze and Maze Tests by Respondent Demographic Group

Respondent Demographic Group	N	Cloze	Maze
By Sex:			
Males	31	.57**	.16
Females	29	.57**	.09
By Race:			
White	27	.52**	.20
Black	33	.57**	.20
By Age:			
Over 35	28	.53**	-.10
Under 35	32	.55**	.45*
TOTAL	60	.53**	.17

*p = .05.

**p = .01.

The correlation between Cloze and Maze is low-positive; the correlation between DSST and SCT is somewhat higher. Each of these consistently related to some other tests. Tables 2.5 on page 20 and 2.6 on page 21 list the intercorrelations of these tests with the other tests. DSST seems to relate consistently to the performance syndrome; the SCT to the verbal syndrome.

These two then seemed like our most promising measures. However, further analyses of the DSST and its relationship to reported anxiety suggested that it was highly influenced by immediate motivational factors. Therefore we sought other non-verbal measures. The Block Design Test was very reliably related to criteria across groups, but was unfeasible for survey work; therefore we turned in the Detroit Survey to a substitute for it. The Digit Span Test was also promising, so we looked for parallel measures of attention that would be related to more complex skills.

We were disappointed in the Miner Vocabulary Test -- its range seemed too limited to produce any results at all. Our thoughts turned to other vocabulary tests.

Each of these considerations led us to design new methods to use along with the most promising ones from Jackson II for evaluating intelligence in another survey.

Our primary effort of the research was to test the feasibility and validity of a survey administration of intelligence measures. The data indicated that a number of scales could be used in an interview setting. The patterns of correlation are similar to those that might be predicted by psychodiagnosticians. However, problems did occur in the interview assessment which suggested further study for the next methodological exploration.

Analyses of data for different interviewers in the Jackson II Study revealed a potential for a considerable interviewer effect on test anxiety -- a variable that could affect both intelligence performance and motivational assessments. Different interviewers seemed to produce vastly different correlations between test anxiety and performance. Such a factor could

Table 2.5

Correlation (Tau Beta) Between Digit Symbol Substitution Test and All Other Tests

	Total Population N=50	Males N=26	Females N=24	Whites N=26	Blacks N=24	Over 35 N=25	Under 35 N=25
Digit Span	.479**	.432*	.585**	.565**	.389	.467*	.483*
Vocabulary	.123	.017	.253	.036	.164	.189	-.029
Blocks	.411**	.680**	.274	.541**	.163	.504**	.351
Information	.236	.402*	.235	.292	.179	.089	.290
Sentence Completion	.280	.357	.276	.371	.161	-.012	.522**
Cloze	.364*	.429*	.328	.399*	.337	.382	.281
Similarity	.344	.381*	.290	.377	.305	.307	.294
Maze	.251	.324	.254	.366	.172	.035	.518**

*p = .05.

**p = .01.

Table 2.6

Correlation (Tau Beta) Between Sentence Completion Test and All Other Tests

	Total Population N=50	Males N=26	Females N=24	Whites N=26	Blacks N=24	Over 35 N=25	Under 35 N=25
Digit Symbol	.280	.357	.276	.371	.161	-.012	.522**
Digit Span	.198	.163	.399*	.292	.093	.032	.324
Vocabulary	.105	.325	-.245	.150	.046	.224	-.070
Blocks	.310*	.383*	.279	.311	.236	.108	.482*
Information	.412**	.471*	.391	.440*	.375	.313	.498**
Cloze	.532**	.571**	.518**	.517**	.573**	.534**	.547**
Similarity	.347*	.371	.363	.324	.383	.239	.415*
Maze	.170	.163	.092	.205	.196	-.102	.469*

*$p \leq .05$.

**$p \leq .01$.

completely undermine assessment efforts. Consequently, we began thinking seriously of studying a critical interviewer effect -- one in which we could have some control -- race of the interviewer, especially in interaction with the race of the respondent.

Where We Stood Prior to the Detroit Survey

Having explored methods of measuring intelligence and motivation in two survey settings and having obtained some meaningful results, we began to feel optimistic about their feasibility but still concerned about how to evaluate validity. In a survey interview the respondents seemed to accept novel methods, but what did they mean?

Intelligence measures had a great deal of internal consistence but we lacked any way to compare them to standards outside of the interview setting. We needed a more heterogeneous population to test the correlation between measures and education, the characteristic against which these tests were validated to begin with. Furthermore, we were alerted to the fact that the race of the interviewer might affect the validity of the assessment. These results clearly called for a cross-sectional survey with some control on race of interviewer and respondent.

The responses to the motivation measures had little internal consistency but did suggest the following:

1. Open-ended assessment techniques, such as telling stories in response to pictures, are susceptible to styles of presentation but in fact are promising. Furthermore, if the story-telling assessment of achievement motivation has the most currency with social psychologists, it should be included in order to get some data on the validity of new procedures.

2. Objective measures do not correlate well with one another, a fact that forces us to examine the psychological significance of the contents of achievement items. A multi-faceted approach to the concept of achievement goals seems called for. A larger sample would permit a more careful analysis of the dimensions of the phenomenology of achievement motivation.

3. The behavioral criteria for evaluating the value of achievement motivation measures are complicated, but feasible. They bear repeating in a more heterogeneous context and are worth considering as major criteria against which to establish validity of new procedures.

The Detroit Survey: Overall Procedure

Results from the Jackson II Study suggested new measures of intelligence and motivation, and ways of reworking old measures. The cross-sectional survey design of the Detroit Study, with its larger and wider spectrum of socio-economic sampling, allowed us to validate measures against a social characteristic such as education, as well as repeat the validity procedures used in Jackson II. This was particularly important for the intelligence measures. A further virtue in the Detroit survey was the opportunity to do a careful test of the effect of race of interviewer -- a factor we thought so critical in assessment of motivation and intelligence measures that it needed systematic study. Half of the respondents were to be black; half of the respondents white. The same was true of the interviewers. A detailed description of this study, its design, and strategy appears below.

The following highlight the special features of the survey:

1. Two significant new measures of intelligence were introduced, the Raven Progressive Matrices and the Picture Order Test (see Appendix 2 for descriptions of both).

2. Some important open-ended procedures for measuring achievement motivation were introduced, especially one that had proved successful for a similar population. We were especially interested in getting at assessment of fear of success in women and blacks.

3. The behavioral measures used as validity criteria in Jackson II were sharpened considerably.

4. Anxiety measures were introduced for testing the interviewers' effects on reported anxiety, and for some theoretically derived predictions of achievement performance. New experimental procedures in the way anxiety is reported to the interviewer were also included. All of these new

factors in motivational assessment will be discussed below. With this overview of the Detroit study, let us proceed with a more detailed description of it.

The Sample

A probability sample of 1,027 households within the city of Detroit was contacted by six black and eight white Survey Research Center interviewers. The addresses sampled were from primarily black, white, and mixed neighborhoods and were designed to yield a cross-section of each race group. (Mexicans, Orientals, and others not identifiable as black or white were eliminated from the sample.) Within each household, a random selection table was used to select the respondent from all residents between 18 and 49. A response rate of 71% yielded the following:

365	Interviews
407	Households with no eligible respondents (i.e., between 18 and 49 years old)
89	Unoccupied addresses
172	Refusals or respondent absent
34	Addresses not contacted

The interviews were stratified into four equal groups by race and sex of respondent: 90 black males; 96 black females; 89 white males; and 90 white females. Approximately half of each group was interviewed by blacks, half by whites.

Measures

At the beginning of the interview, each respondent was asked about his marital status, his occupation, and the occupation of his spouse and father. Later, the highest grade completed and personal 1969 income for both respondent and spouse were assessed. These data were collected to provide for both demographic controls on the sample and external validity checks on the motivation and intelligence measures.

Seven intelligence measures and several different types of motivation measures were administered in the main part of the interview; these are described fully in the following two sections, along with the methodological consideration attached to each.

CHAPTER THREE
VALIDATING MEASURES OF INTELLIGENCE

At the conclusion of this chapter, we recommend the adoption of the Sentence Completion Test from the Lorge-Thorndike intelligence test as a feasible, reasonably valid assessment of what psychologists have labeled intelligence. We did not expect to recommend just one measure; rather, we expected to find two types of measures, each reflecting a different kind of intelligence commonly considered in the literature on intellectual competence: a verbal mediational facility and a perceptual performance facility. What we found is that one test, the Sentence Completion Test, seems to correlate well with most different kinds of tests of intelligence, well enough to suggest using it singly without going to multiple measurement. When we took a close look at the test, we realized why this is so. The test asks the respondent to supply from a set of alternatives a missing word in a sentence. Certainly, it is a measure of verbal comprehension and learning. In spite of this verbal emphasis, however, the SCT also requires hypothesis-testing and skill in patterning sentences similar to the skills involved in some of the perceptual performance measures we used.

Measures

Seven intelligence measures were included in the Detroit survey. All are short, and feasible in a household interview. According to clusters previously mentioned they are:

Verbal mediational facility:
- Sentence Completion Test (Lorge-Thorndike)
- Ammons Quick Test
- Information (Weschler)

Perceptual performance facility:
- Digit Span (Weschler)
- Raven Progressive Matrices
- Digit Symbol Substitution (Weschler)
- Picture Order, Central and Incidental

For a complete description of each test, see Appendix 2. Sentence Completion and Information, as used in Jackson II, were closely related to our mediational criterion, the Cloze Test. The Ammons was included because it had previously appeared promising in discrimination at the low end of the scale -- a quality not characteristic of the Vocabulary Test used in Jackson II. The standard procedure for the Ammons was revised, and some items were eliminated.

Because the interview was long, not all tests could be given to all respondents. Therefore, a random half received the Ammons and Information Tests; the other half, the Picture Order Test. All other tests were administered to all respondents.

Digit Symbol Substitution and Digit Span correlated well with the attentional performance criterion, Porteus Mazes, in Jackson II. In addition, the Digit Symbol Substitution also correlated highly with the Cloze Test.

Most critically, we introduced two new assessment procedures for intelligence in the Detroit survey: Raven Progressive Matrices and the Picture Order Test. Koh's Block Design Test had proved very promising in Jackson II, but it was clearly unfeasible for general household survey use. This test was replaced by the Progressive Matrices, which has been shown to be highly related to the Koh's Test (Hall, 1957). The Matrices is an easily administered perceptual task of choosing one pattern to complete a larger design. The Picture Order Test was selected because it seemed to be a more complex version of the skill involved in the Digit Span -- attending to the position in which a given picture occurred. It is a measure of attentional facility. Secondarily, it also allows for measurement of incidental learning, another possible factor involved in competence. The work on this test has been very significant in plotting theoretically meaningful developmental shifts in attention (Hagen and Huntsman, 1969). These developmental changes may also be useful in assessing individual differences -- especially those involving brain injury. The use of this measure is the most innovative of all our procedures.

The Detroit survey thus uses a variety of methods of administration of the intelligence tests. These can be characterized as follows:

Table 3.1

Intelligence Measures Used in the Detroit Survey

Test	Type of Presentation	Response Mode
Sentence Completion	Verbal/written	Verbal -- choice of 5 words
Ammons Quick Test	Verbal	Point to one of 4 pictures
Information	Verbal	Verbal (open)
Digit Span	Verbal	Verbal
Raven Progressive Matrices	Graphic	Point to one of 8 patterns
Digit Symbol Substitution	Graphic	Written (timed)
Picture Order Central	Graphic	Point to one of a series of pictures
Picture Order Incidental	Graphic	Match one set of pictures with another

Analysis

We followed three steps in the analysis of data summarizing the different measures of intelligence employed in the survey. First, we intercorrelated all these measures, in hopes of finding the two clusters we had hypothesized -- a verbal facility and a perceptual performance facility, and in hopes of locating the one test in each cluster that accounted for the most variance in the intercorrelations. As we indicated, this clustering was not apparent, so we just asked of the matrix which test was the most effective in the intercorrelations in general, or which test correlated highest with other tests.

Secondly, we looked at the correlations of the various tests with education, recognizing the nature of intelligence assessment used in research and in practice in previous years. Ever since Binet began the standardized assessment of intelligence to predict school success, developers of measures of intelligence have relied explicitly or implicitly on educational success as their main validity criterion. Usually a new test was "validated" against an old one. Therefore, Binet's criterion was the ultimate one. Consequently, we turned to education attained by a respondent as a criterion against which to assess the validity of the measures we used.

Thirdly, we returned again to analysis of the intercorrelations of the various measures of intelligence, this time controlling for various demographic correlates of the intelligence measures. We wanted to make sure that any recommended intelligence test had some variance left over after various demographic factors, especially education, were controlled. We reasoned that if the measures were not reasonably highly interrelated at this point, using a recommended intelligence test would not yield any increase in predictability over just using the level of the respondent's education.

Intercorrelations of Measures of Intelligence

We hypothesized that the eight different measures of intelligence would cluster into two distinct groups corresponding to a hypothetical verbal mediational facility and a hypothetical perceptual performance facility. We expected the Digit Span, Raven Matrices, Picture Order Central, Picture Order Incidental, and Digit Symbol Substitution tests to cluster into the perceptual performance ability cluster, and the Sentence Completion, Ammons, and Information tests to cluster into the verbal mediational facility.

Before proceeding to the intercorrelation matrix, however, we wanted to be sure that none of the measures showed any interviewer bias. We were concerned about the possibility that both black and white respondents would operate differently with a white interviewer than with a black interviewer. Especially of concern was the distinct possibility of an interaction effect: that black respondents would be more anxious and do less well in response to a white interviewer than in response to a black interviewer, and vice versa. If so, this would present some really critical methodological problems for using both population norms and individual data on intelligence. Correction might be so unwieldy as to disallow the usefulness of such measures. In Appendix 4, we discuss our examination of the interviewer effect, not only on intelligence but on motivational measurement also. We conclude that there are no interactions that would particularly contest the use of the scores, unadjusted for the race of the interviewer, for our correlational analyses.

This result is summarized in Table 3.2 below, where we present the mean score on each of the intelligence tests for eight groups: the population, partitioned exhaustively by sex, race, and race of interviewer. The analyses of variance of the effects of these variables on the mean scores are presented in Table 3.3 on page 32. There we see that there is no significant race of interviewer-race of respondent interaction; in fact, there is only one interaction effect altogether that is significant, one that we can perhaps discount since there were so many possible interactions that could have been significant. In the tables we see the pronounced race difference in mean measured intelligence scores, with only the Picture Order Incidental showing no significant difference. There is also a significant race of interviewer effect. Black interviewers generally were eliciting higher mean intelligence scores from all groups. These results can be partly explained by the fact that the black interviewers interviewed a slightly more educated population. Table 3.4 on page 33 shows clearly that the black interviewers' respondents, both black and white, were slightly more educated. However, the effect of the race of the interviewer on intelligence remained

Table 3.2

Mean Scores on Intelligence Measures, for Eight Population Groups

	White Interviewer				Black Interviewer			
Respondents' (Race)	White		Black		White		Black	
Respondents' (Sex)	Males	Females	Males	Females	Males	Females	Males	Females
Measure								
Digit Span	5.15	5.16	4.60	4.58	5.87	5.37	4.93	5.04
Sentence Completion	11.19	11.07	9.46	10.28	11.87	11.34	10.90	10.70
Raven Matrices	6.96	6.18	4.33	4.23	7.74	6.23	5.83	4.89
Ammons	39.40	40.63	35.13	36.35	40.22	39.04	37.42	34.57
Information	9.00	8.30	6.33	5.80	9.50	7.31	7.38	6.72
Picture Order Central	8.52	8.50	6.97	7.52	8.42	9.08	8.05	8.23
Picture Order Incidental	2.92	1.72	2.52	2.26	3.37	3.25	3.60	3.17
Digit Symbol Substitution	30.77	35.18	25.47	29.65	34.55	34.80	26.78	28.40

Table 3.3

F - Ratios from the Analyses of Variance on Eight Intelligence Measures: Race of Interviewer x Race x Sex of Respondent

F-Ratios	Inter. Race (A)	Resp. Race (B)	Resp. Sex (C)	A x B	A x C	B x C	AxBxC
Digit Span	7.11**	11.90**	0.13	0.04	0.30	0.55	0.93
Sentence Completion	10.62**	21.13**	0.10	0.99	2.57	1.88	0.46
Raven	11.12**	71.34**	9.86**	1.99	2.79	1.60	0.02
Ammons	0.07	43.90**	0.07	0.22	7.06**	0.45	0.46
Information	0.31	13.37**	7.41**	1.67	0.98	1.57	0.48
Picture Order Central	3.48*	8.31**	1.13	1.09	0.05	0.07	0.65
Picture Order Incidental	15.96**	0.94	4.87*	0.01	0.53	0.92	1.46
Digit Symbol Substitution	0.71	42.24**	8.30**	0.73	2.98	0.01	0.17

*p ≤ .05.

**p ≤ .01.

Table 3.4
Mean Education Attained, by Interviewer and Respondent Race

		Respondent Race	
		White	Black
Interviewer Race	White	11.88	10.71
	Black	12.73	11.27

even when education was used as a covariate in an analysis of covariance. (See Appendix 4 for this analysis.) Therefore, we have a definite interviewer race effect, an effect we did not expect. Black interviewers elicited higher scores on intelligence tests on the average from everyone. Further discussion of this phenomena is found in Appendix 4.

For our current purposes, then, we felt assured that the measures were not contaminated by any interaction of race of respondent with the race of interviewer. We then proceeded to look at the intercorrelations of the eight measures of intelligence with each other. These results are found in Table 3.5 on page 35. Note that four cells of the matrix are blank; these involve the Picture Order Test, which was given to only half the sample, and the Ammons and Information Tests, which were substituted for it in the other half of the population.

In examining Table 3.5, we can see that the separate clustering we were hypothesizing did not emerge. All the tests seem to be equally intercorrelated, suggesting that we are picking up a very general factor of intellectual competence. We therefore abandoned our search for two different types of intelligence measures, and began to ask which one of the measures stood out in accounting for the most variance.

We computed the intercorrelation matrix separately for blacks and whites, anticipating that the significance of the measures might be slightly different for the two racial groups. What is impressive is that the intercorrelations for the two different groups are in fact very similar, except perhaps those generated by the Picture Order Incidental Test.

The answer to this question is very clear, and confirms the results from the Jackson II Study. The set of correlations generated by the Sentence Completion Test is higher than that generated by any other test, across the whole matrix. In fact, only in one instance does another test, the Digit Symbol Substitution Test, show any higher intercorrelation. This was our first clue that the Sentence Completion Test was the one we would recommend for OEO use.

We were concerned about the potential bias that might arise from taking such a verbal task as the SCT as a standard for testing intelligence in a

Table 3.5

Intercorrelations of Intelligence Measures, by Respondent Race

Measure	DS	SC	RAV	AMM	INFO	P.O.C.	P.O.I.	DSS
Digit Span		.46	.41	.39	.30	.38	.22	.47
(DS)		.40	.30	.21	.23	.33	.05*	.36
Sentence Completion	177		.54	.55	.45	.42	.22	.49
(SC)	188		.43	.62	.52	.48	.28	.45
Raven	177	177		.45	.46	.30	.08*	.43
(RAV)	188	188		.44	.34	.40	.24	.41
Ammons	85	85	85		.46			.34
(AMM)	85	85	85		.46			.44
Information	85	85	85	85				.39
(INFO)	83	83	83	83				.24
Picture Order								
Central	90	90	90	0	0		.07*	.40
(P.O.C.)	102	102	102	0	0		.10*	.39
Picture Order								
Incidental	89	89	89	0	0	89		.09*
(P.O.I.)	102	102	102	0	0	102		.27
Digit Symbol								
Substitution	177	177	177	85	85	90	89	
(DSS)	188	188	188	83	83	102	102	

Note:

Correlations are above diagonal, number in sample below diagonal.

Whites are on top row, blacks below.

All correlations are significant at at least $p \leq .05$, except those marked by an asterisk.

heterogeneous population. Therefore, we wanted to explore as carefully as possible whether any of the intercorrelation results found in Table 3.5 were in fact determined only by certain population groups, especially those from higher socio-economic backgrounds. Would the Sentence Completion Test generate as strikingly strong correlations for people from disadvantaged backgrounds as it would for people from advantaged backgrounds?

We tried to answer this question by recomputing the intercorrelation matrices found in Table 3.5 for people coming from four different socio-economic backgrounds, as measured by the socio-economic status of the father (reported by the respondent). For this status distinction we used Duncan's prestige coding of the father's occupation, dividing the respondents into three groups as nearly equal in size as possible. We computed the intercorrelations separately for those people who reported their father as being absent from the home while they were growing up. These four different sets of intercorrelations of the eight tests of intelligence are reported in Table 3.6 on page 37. The table clearly indicates that the superiority of the SCT above all the other tests is apparent in each matrix. If we were to judge how useful the tests were from the intercorrelations alone, we would clearly state that the Sentence Completion Test is a superior test for people from different backgrounds.

One exciting additional finding which emerged from examination of these four different sets of intercorrelation matrices is worth noting. The correlations generated by the Picture Order Incidental Test are markedly higher in the two lower status groups than in the other two groups. This result and others concerning the Picture Order Incidental Test will be discussed further in Appendix 6.

In summary, from our examination of the intercorrelations of the tests with each other, we conclude that the Sentence Completion Test seems to be the best single measure of general intelligence.

Table 3.6

Intercorrelations of Intelligence Measures, by Socio-Economic Background Groups

Test	SC	RAV	AMM	INFO	P.O.C.	P.O.I.	DSS
No Father in Home (N = 78)							
Digit Span	.37**	.30**	.26	.13	.35**	.38**	.44**
Sentence Completion		.42**	.62**	.40*	.45**	.38**	.44**
Raven			.39*	.12	.40**	.45**	.46**
Ammons				.44**	---	---	.33*
Information					---	---	.09
Picture Order Central						.17	.44*
Picture Order Incidental							.63**
Low Socio-Economic Background (N = 79)							
Digit Span	.50**	.39**	.30*	.18	.27	.07	.31**
Sentence Completion		.58**	.67**	.49**	.55**	.27	.54**
Raven			.56**	.38*	.33*	.24	.50**
Ammons				.52**	---	---	.60**
Information					---	---	.35*
Picture Order Central						.07	.34*
Picture Order Incidental							.31*
Medium Socio-Economic Background (N = 128)							
Digit Span	.47**	.40**	.35**	.29*	.36**	.09	.48**
Sentence Completion		.46**	.60**	.59**	.22	.09	.50**
Raven			.52**	.54**	.28*	-.09	.40**
Ammons				.49**	---	---	.46**
Information					---	---	.46**
Picture Order Central						.16	.26*
Picture Order Incidental							-.08
High Socio-Economic Background (N = 80)							
Digit Span	.34**	.34**	.28	.52**	.32	-.05	.32**
Sentence Completion		.51**	.49**	.49**	.43**	.22	.33**
Raven			.29	.56**	.30	-.04	.37**
Ammons				.50**	---	---	.22
Information					---	---	.36*
Picture Order Central						-.10	.34*
Picture Order Incidental							-.17

*$p \leq .05$.
**$p \leq .01$.

Correlations of Intelligence Measures with Education

When Binet was constructing the first standardized intelligence test in France, his goal was to detect those children who would not do well in school. Although he used age grading to diagnose the usefulness of a particular item, the validity of the test as a whole depended on the practical usefulness it had for detecting children who would fail in school. Similar validity studies have been done since Binet's original work. Following Binet's standardization and ones subsequent to it, researchers in intelligence used the standardized tests themselves as criteria by which to evaluate new tests. Therefore, the educational achievement of children served as the ultimate criterion for the validity of most new tests.

We adopted a similar criterion for our measure. We could not measure success or failure in school directly, but we could measure how far in school the respondent went; this is certainly a multidimensionally determined measure, but one that could depend on intellectual competence, to some extent, as in Binet's criterion. We related the educational level attained by the respondent to the measures of intelligence; the results are summarized in Table 3.7 below. In that table, we can see that only one other test, the Digit Symbol Substitution Test, correlated higher over-all with education than the test we are going to recommend, the Sentence Completion Test.

Table 3.7

Correlations of Education with Intelligence Measures

Intelligence Measure	r
Digit Span	.41**
Sentence Completion	.57**
Raven	.50**
Ammons	.51**
Information	.35**
Picture Order Central	.43**
Picture Order Incidental	.15*
Digit Symbol Substitution	.62**

*p $\leq$ .05.

**p $\leq$ .01.

Table 3.8 below shows the correlations of the intelligence measures with education, separated by racial group and by father's socio-economic rating.

Table 3.8

Correlations of Education with Intelligence Measures, by Race and Socio-Economic Background Groups

	Race			Socio-Economic Background		
Measure	Whites	Blacks	Father Absent	Low SES Background	Medium SES Background	High SES Background
Digit Span	.42*	.35*	.47*	.34*	.30*	.40*
Sentence Completion	.51*	.58*	.58*	.61*	.50*	.45*
Raven	.48*	.43*	.39*	.56*	.39*	.41*
Ammons	.51*	.47*	.44*	.67*	.37*	.32*
Information	.48*	.44*	.41*	.43*	.49*	.52*
Picture Order Central	.38*	.42*	.46*	.57*	.15	.16
Picture Order Incidental	.03	.29*	.50*	.14	-.12	.05
Digit Symbol Substitution	.59*	.59*	.52*	.66*	.56*	.47*
(N)	(177)	~~(188)~~	(78)	(79)	(128)	(80)

*$p \le .01$.

These were computed in order to see whether the high correlations between the SCT and education found over the whole group might be due not to the common factor of intelligence, but rather to a third variable, culturally-determined opportunities for education and intellectual development. It may be that social class, or race, determines both educational attainment level and Sentence Completion Test scores. The data in Table 3.7 show quite reassuringly that, even grouping on race and father's socio-economic status as gauges of culturally determined educational opportunities, the correlations between SCT scores and education remain rather high and statistically significant.

We should note that for the black group, the correlation between Picture Order Incidental and education is significantly higher than it is for the white respondents. There thus seems to be mounting evidence that this measure may be reflecting a very special type of facility for some minority-group, lower-status respondents.

Correlating the Sentence Completion Test with Other Tests, Partialing Out Demographic Correlates

Since the correlations of the Sentence Completion Test with education were so consistent and striking, we were concerned that we might be measuring only educational attainment, rather than any surplus facility called intelligence. Therefore, we re-examined the correlations of the SCT with the other measures of intelligence, this time partialing out the effects of education and other social factors that were correlated with all the intelligence measures (see Table 3.9 on next page). The correlations, although smaller than those reported in the table of raw correlations, are still significant and reasonably high, suggesting that the Sentence Completion Test especially is more than a proxy for educational attainment or other demographic variance.

Table 3.9

Intercorrelations of Intelligence Measures, with Education, Age, Race, Sex, Interviewer Race, and Region Partialed Out

Test	SC	Raven	Amm.	Info.	P.O.C.	P.O.I.	DSS
Digit Span	.28**	.24**	.12	-.11	.26**	.06	.29**
Sentence Completion		.33**	.47**	.38**	.30**	.18*	.23**
Raven			.28**	.22**	.19*	.05	.22**
Ammons				.33**	---	---	.19*
Information					---	---	.15*
Picture Order Central						.02	.12
Picture Order Incidental							.17*

*p ≤ .05.

**p ≤ .01.

A Measure of Intelligence to be Used in Surveys: the Sentence Completion Test

We are thus suggesting that a thirteen item test, reproduced in Table 3.10 on the following pages, from the verbal part of the Lorge-Thorndike Intelligence Test (1950) could be used in sample surveys. It should not take more than an average of five minutes to administer. Each item is a sentence in which a word is missing. The respondent is given five words from which he is to select the one that makes the best, truest, most sensible sentence.

Table 3.10

Sentence Completion Test

Here's another type of test. Turn to page _ in your booklet. I'll read each of these sentences. In each one there's a word left out. After I read the sentence, you pick one of the words under the sentence to complete it. Say the word that you think makes the best, truest, most sensible complete sentence. If you're not sure of the answer, take a guess.

For example: "Lemons are sour, but sugar is "Blank." Which word would you pick? (READ ALL ANSWERS, WORDS AND NUMBERS.) "One, bitter; two, white; three, fattening; four, sweet." (READ ALL ANSWERS, EVEN IF INTERRUPTED.)

IF R PICKS "FOUR, SWEET," SAY "O.K., let's try some more."

For the rest, I won't tell you if you're right or wrong. Just give me your answer.

IF R GIVES THE WRONG ANSWER, say, "No, it's number four, sweet. Lemons are sour, but sugar is sweet. Now, let's try some more."

DO NOT ACCEPT "Don't know" RESPONSES OR REFUSALS TO GUESS.
DO NOT REINFORCE ANY ANSWERS.
USE THE TRANSITION STATEMENTS.
CIRCLE THE ANSWER GIVEN.

SC1. We see ____________ only at night.

1) Children 2) Plants 3) Stars 4) Houses 5) Trees

Here's the next one SC2. Not every cloud gives ____________.

1) Weather 2) Shade 3) Sky 4) Climate 5) Rain

O.K. SC3. In the spring the buds form on the branches of the ____________.

1) Trees 2) Rivers 3) Bugs 4) Leaves 5) Animals

Table 3.10 -- Continued

Mm Hmm	SC4.	There is an old ____________, "An apple a day keeps the doctor away."	1) Talk	2) Saying	3) Reader	4) Book	5) Man
The next question	SC5.	The ragged ____________ may prove a good horse.	1) Puppy	2) Child	3) Calf	4) Lamb	5 Colt
Let's try the next one	SC6.	The important thing is not so much that every child should be taught as that every child should be given the wish to ____________.	1) Learn	2) Play	3) Hope	4) Reject	5) Teach
OK	SC7.	The person who ____________ another must make good the damages.	1) Reforms	2) Improves	3) Instructs	4) Injures	5) Delights
All right	SC8.	False facts are highly __________ to the progress of science.	1) Injurious	2) Necessary	3) Devoted	4) Useful	5) Instrumental
Number 9 is	SC9.	It is better that ten guilty persons ____________, than that one innocent suffer.	1) Suffer	2) Escape	3) Capture	4) Starve	5) Repent
Mm Hmm	SC10.	The winds and the waves are always on the side of the ablest ____________.	1) Soldiers	2) Statesmen	3) Navigators	4) Students	5) Weathers

Table 3.10 -- Continued

O.K.	SC11.	The vanquished never yet spoke ____________ of the conqueror.				
		1) Ill	2) Well	3) Little	4) Nastily	5) Often
Here's the next one	SC12.	Think long when you may ____________ only once.				
		1) Abstain	2) Live	3) Die	4) Decide	5) Eat
O.K.	SC13.	The coward threatens only when he is ____________.				
		1) Afraid	2) Surrounded	3) Safe	4) Conquered	5) Happy

We had originally used a fifteen-item test, found on page 18 of the interview (in Appendix 1). However, two of the items (numbers 1 and 15) had extremely low average inter-item correlations; hence we do not recommend them for further use. One of these is still part of the test as an initial "warm-up" item, but should not be counted in the score.

This test discriminates best at the lower end of the distribution (see Figure 3.1 on page 45). The internal consistency of the instrument (correlation of even-item scale to odd-item scale), .46, is relatively low, but that should not be of too much concern since the test is very short. There are a number of component skills involved in answering a given item of the test successfully (e.g., vocabulary, grammar awareness, logical reasoning). The different items probably tap these skills in different proportions. We would guess that the test-retest reliability of the Sentence Completion Test should be much higher.

Use of the Test

One major warning should accompany the recommendation of the Sentence Completion Test for further use. There was a definite effect of race of interviewer on the scores, with black interviewers' respondents scoring higher. Because interviewer race did not interact with any other variables to produce complex effects, and because we were depending almost entirely on correlational analyses, we did not find it necessary to remove this effect.

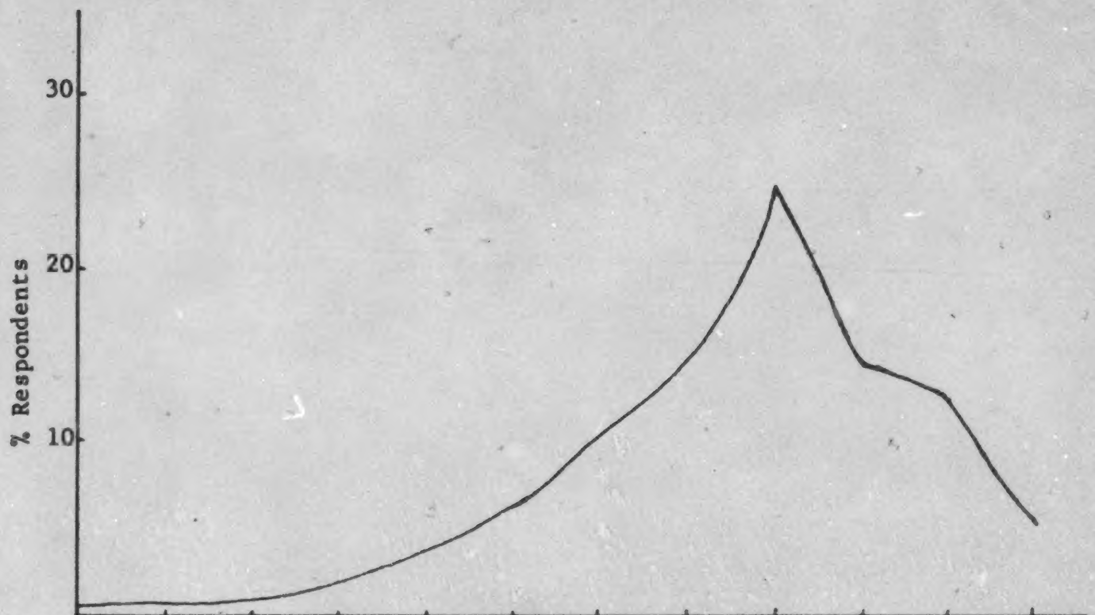
However, we would caution future users of the SCT to be alert for possible interviewer race effects. Such an effect can be easily tested for and removed if interviewer race is not confounded with respondent race and if no interactions occur. This would be recommended for use in any analyses involving individual prediction, as would a correction for a mean difference due to the race of the respondent.

Thus the SCT is easy and quick to administer, understand and score, long enough to give a stable estimate, and broad in scope, tapping several different facets of intelligence. It is a measure of "standard" intelligence

Figure 3.1

Distribution of the Scores

Obtained on the Sentence Completion Test



which relates strongly to but is more than attained education. Although it may be affected by the race of the administrator, it otherwise seems ideal for use in large household surveys of varied populations.

CHAPTER FOUR

METHODS OF ASSESSING ACHIEVEMENT MOTIVATION IN A SURVEY: THEORETICAL BACKGROUND AND DESIGN OF STUDY

Since McClelland's early work on achievement motivation (1953), research on the meaning of the achievement motive has forced the motivation theorist to broaden the concept of the achievement motive into several distinct components. The idea of a unitary "achievement motive" was first questioned in the Atkinson (1964) theory, in which motive to succeed and motive to avoid failure interact to produce a resultant motive. Since then, Atkinson has separated expectancies and incentives from motives; Veroff (1965) and Birch and Veroff (1966) have justified thinking about additional, dispositional expectancies; Horner (1968) has shown the importance of a motive to avoid success, especially in women; Veroff (1969) has distinguished motivations for autonomous standards of excellence from those for more social standards; and Raynor (1969) has suggested that future orientation is a critical factor in determining persistent achievement performance.

Such developments can lead to considering achievement motivation as an "umbrella" concept, embracing many distinct components. The breadth and variety of these components has been explored in the past ten years of research on achievement; how they combine and weight to predict different behaviors has not been so well explored.

In this connection, it is revealing to consider two of the behaviors that have figured heavily in achievement motivation research over the past two decades -- fantasy behavior from which a measure of achievement motivation is derived, and risk-taking behavior, which has become one of the most critical criteria for evidence of resultant achievement motivation since Atkinson first developed his thinking on risk-taking in 1957.

Fantasy measures of achievement motivation have become the standard measure of achievement motivation, ones by which more innovative assessment techniques are evaluated. Fantasy measures were developed originally by contrasting people in "achievement situations" (e.g., ego-involvement about doing well on a test) to those not in such situations. As such, projective imagery, usually in response to selected TAT pictures and carefully coded by a highly reliable system (Atkinson, 1958), represents both a result of and a synthesis of many of the separate achievement components. Whether a person in a story is fearful or hopeful, whether negative consequences occur if he succeeds, whether he has highly internalized standards of excellence, or uses those defined by society, are not differentiated in the over-all score. Thus, the achievement motivation score represents a composite of potential components of achievement orientation. Indeed, the success of the measure may depend on its being an undifferentiated potpourri of various achievement attitudes.

Now let's look at risk-taking behavior. Atkinson's theory of risk-taking (1957) hypothesizes that people with high achievement motivation like to be challenged by the work they do or the games they play. They like to succeed but only when there is some chance of failure. Specifically, Atkinson theorizes that they like to operate best at a 50% probability of success. Since Atkinson's initial statement, a number of empirical studies have supported this theory. More recently, research by Heckhauson (1967) has suggested that the preferred risk for high achievers may be at a more difficult level than was originally thought, as low as 30% chance of success. Nevertheless, the generalization that moderate levels of risk are preferred to extremes by highly achievement motivated people still seems to hold. Operationally, this has usually been taken to mean a selection of moderate risk situations, and better performance or persistence in such situations by highly motivated subjects. When one closely examines moderate risk taking, however, it would be hard to pinpoint any one of the aforementioned components as being dominant in all situations across the board. Fear of failure, fear of success, future orientation, autonomous standards, social comparison standards can all be involved. Certainly in any given situation, one of these may predominate. And indeed one can experimentally manipulate risk-taking behavior by emphasizing one or another of these components (Heckhauson, 1967). But, across the board, risk-taking behavior seems to be a catchall behavior, reflecting the operation of many components of achievement orientation at the same time.

Thus, projective achievement imagery responses and moderate risk-taking are two types of achievement behaviors which represent end-products of many different components of achievement motivation. As such, they are useful behavior indices of motivation to ferret out potential refinement in assessment and theoretical analysis. We will make use of this conclusion about these behaviors in our own procedure in this study.

The fragmentation of the once-unitary concept of achievement motive has been partially the result of work with new subject populations. Most of the early theory of achievement motivation was the result of data collected from white college males. When verification on other populations was attempted, the theory or the validity of known assessment procedures often were not supported. This research, plus related sociological research and observation, point to two population divisions important to work in achievement motivation: division by sex and by race in the United States.

A large amount of psychological and sociological research and theory supports separate examinations of men and women. In their 1962 national survey of the dimensions of subjective adjustment, Veroff, Feld, and Gurin found it necessary to do separate factor analyses for men and women and arrived at considerably different factor structures for each sex. Horner (1968) has cogently argued, from data and theory, that the motivations primarily effective in understanding women's achievement preferences and performances are quite different from those of men. She contends that fear of success may be a strong inhibiting factor for American women, one that interferes with their assertiveness in competition with one another as well as with men. Elsewhere, Veroff, Wilcox, and Atkinson (1953), French and Lesser (1964), Pierce and Bowman (1960), Angilini (1955), Lesser, Krawitz and Packard (1963), and most recently Fontana (1970) have produced research evidence that the relationship of fantasy measures of achievement motivation to social situations and to types of performance is different for women and men.

The findings of Dornbusch (1966), Mead (1935), Kluckhohn (1953), Mannes (1963), and Parsons (1949) are just a few examples of sociological statements about how American society induces a different type of achievement orientation in women than in men. The crux of the sociological position is that a woman's achievement orientation finds expression in the home, chiefly in response to the social demands of her family, while a man's achievement orientation is in the business and professional world outside of the home, chiefly in response to his own internal strivings. Indeed, Veroff and Feld (1970) have found that men and women with high achievement motivation approach both marriage and parenthood in very different styles, corresponding to the sociological distinctions we have just mentioned.

Common sense and sociological observations dictate that race is a factor with as much potential impact on achievement orientation as sex. The barriers to achievement that have existed for blacks in our society need no documentation here. Such barriers may create for blacks a peculiar social environment with regard to the development of achievement motivation, the meaning of moderate risk-taking, and the whole meaning of success and failure. Katz (1967) and Veroff and Peele (1969) have pointed to the overly rigid, unrealistically high standards of excellence sometimes developed by blacks which may interfere with their own expectations of success. Also particularly striking is the possibility that the same fear of success orientation that pervades women's achievement orientation might also exist for many blacks in our society.

In a group for whom achievement is so difficult, if an individual does achieve on middle-class terms, there is a possible psychological disruption of the person's own sense of identity with the group of his origin. Such conflicts about the meaning of success, coupled with very low expectations of success, may produce in blacks an entirely different psychological structure about achievement. Furthermore, the family pattern of most blacks raised in urban society is strikingly different from most white family patterns, which may affect how both men and women from black groups structure their environment for achievement. If the role of provider is thrown on the shoulders of women in urban black society, and men are barred from dignified work, black men and women may get a very different perspective in the learning of achievement orientation than their white counterparts. Finally, the school environments for blacks and whites are critically different. Generally, the ghetto school reinforces the black child's low evaluation of his own competence and hope of success.

Although race and sex would at this point seem to be the most powerful demographic influences upon achievement motivation, many others also deserve consideration. For those of different socio-economic status, intelligence, religion, or type of job, the meaning of achievement might also vary. Clearly, the possibilities are endless.

Thus, the "achievement motive" has grown from a unitary concept with one meaning for all population groups to an umbrella concept covering many diverse components, each of which might carry different meanings for different population groups. Implications for the measurement of achievement motivation are immense. No longer can one hope to make accurate predictions about achievement behaviors on the basis of one motivation score derived in the same way for all.

More specifically, this new conception has three clear implications for measurement:

1. Individual components of achievement orientation must be measured. These likely include motive to succeed, fear of failure, fear of success, autonomous achievement, expectancies, and so on.

2. A component may need to be measured differently in different segments of the population.

3. The components may combine and weight differently for various achievement contexts (job, school, leisure, etc.) or for different population groups.

Clearly, the ultimate extension of these three conditions would result in no prediction at all, for one can infinitely fragment components, populations, and behaviors. The increase in accuracy gained by such fragmentation (obviously, the best "prediction" is after the fact), must be balanced against the decreased practical and theoretical value.

Plan of the Interview

In keeping with the above dictates for measurement, the Detroit interview included many varied achievement measures. These included several new methods, plus revisions of some traditional ones; they are listed below according to the type of response required.

1. Story-telling in response to pictorial and verbal cues. This task was placed near the beginning of the interview with a special fill-in-the-blank "warm-up" exercise preceeding it, to insure adequate responses.* The four story-telling items appear on pages 7 to 10 of the interview (Appendix 1). The first two are based on the four pictures which produced the most interesting achievement responses in the 1960 national survey: a blue collar work scene (picture 1) and a white collar work scene (picture 2) for men; a career setting (picture 3) and a homemaking setting (picture 4) for women. The third item, a verbal cue ("Ted/Joan never hesitates to give his/her opinion."), was meant to elicit power concerns, but was also intended as a warm-up for the fourth item. The last item ("After first term final exams, Phil/Jill finds himself/herself at the top of his/her class at medical school.") was the item most diagnostic of fear of success in women in Horner's research (1968).

Appendix 2 describes in full the coding of these four items for components of achievement motivation. Briefly, "standard" achievement motivation (from the 1958 Atkinson manual), power motivation (Veroff, 1958), type of outcome (positive, negative, or no outcome), instrumentality, and blocks to achievement were coded on all four items. In addition, mastery, social comparison, social approval, autonomy, future orientation, personal efficacy, fear of failure (Birney, et al., 1969), and fear of success (Horner, 1968) were coded on items 1, 2, and 4.

---

*The technique apparently was successful: only 4 of the 365 protocols, or 1%, were not codable, compared to 15% in a 1960 national survey (Veroff, et al., 1960).

2. Shorter open-ended items of completing a sentence about oneself. These items were: "In the next ten years I'm going to...", and "Compared to other men/women my age, I...", from page 8 of the interview. Background for this technique came from the Michigan State University studies predicting economic behavior in Columbian farmers (Rogers and Neill, 1966). The items tap future orientation and social comparison and were coded for "standard" achievement motivation.

3. Self-report in forced choices among either two or three alternatives. Many of these items were adopted from the Jackson II Study andother previous work in motivation, with two major revisions: a forced choice format was used to avoid the response bias that was confounding the responses to the agree-disagree format; and more concrete situations were used as the context for the items to increase the personal relevance and understanding to the respondents. See pages 9-12, 21, and 26 of the interview.

4. Anxiety measures. These techniques received special methodological attention. First, seven items from the Mandler-Sarason Test Anxiety Questionnaire were selected from the twelve used in Jackson. The test had a very high internal consistency, and so could be shortened without appreciably lowering reliability. Second, the items were changed from agree-disagree format to a three-point check-scale to avoid response bias and increase discrimination. Next, an "interview experience" scale (pages 26-27) parallel to the more usual "past test experience" scale was constructed to tap anxiety in the interview as contrasted to more chronic anxiety. This distinction plays a major part in assessment of the interviewer-race effect in case respondents report different anxiety responses to a black than to a white interviewer not only about past performance but also about present performance. Finally, two methods of administration, "public" (respondent gives oral answers to interviewer) and "private" (respondent marks his own answers, which the interviewer does not see) were adopted with the "past test experience" scale, as a way to see if there was any indication that respondents make an "appropriate" anxiety response publicly that they don't do privately. Results of this manipulation are reported in Appendix 5;

they show no differences. All "interview experience" scales were administered in the private condition.

5. Choice of moderate risk in actual performance settings. After each of five intelligence test measures, the respondents were given an opportunity to choose a moderately difficult intelligence item (see pages 15-16, 20, 23, 29, 31 of the interview Appendix 1) in the following way.* The interviewer said, after the test was completed: "I've scored your performance so far on this test. Let's try just one more of these items. Which one would you like to try?

a. One that should be easy for you?

b. One that should be hard for you?

c. Or something in between?

If the respondent chose c, he was further asked: "Like most of the ones you got right, or a little harder than most of the ones you got right?" The last response was coded high on moderate risk-taking. This measure was derived from the Atkinson (1957) theory that people with high achievement orientation prefer to operate under a moderate, or 50% chance of success.

6. Measures derived from the Digit Symbol Substitution Tests (pages 24-25 of the interview). Another assessment of moderate risk in the interview came in connection with the DSST, in which the task was to write in symbols for numbers according to a prescribed code. Each respondent was given the Digit Symbol task for one minute. He was then offered visual examples of forms that were either shorter or just a little longer than his initial base rate performance, and was asked to choose one form. If he chose one just a little longer than his base rate, he was classified as someone choosing a moderate risk; if he chose a shorter one, he was classified as someone not choosing moderate risk. This classification comprised the Digit Symbol Substitution Moderate Preference.

---

*This procedure was slightly different for Digit Span choice (see pages 15-16 of the Interview Appendix).

A second measure derived from the DSST is the differential between a person's performance on a task in which he expects to be able to do well and his performance on an obviously impossible task, the number of symbols to be transcribed being much longer than he had previously been able to do. The measure, Differential DSS Performance, is formed by subtracting rate on Form L (long) from rate on Form M (moderate). (See Appendix 2 for a complete explanation of these different forms.) Thus the respondent serves as his own control, thereby minimizing the effect of individual differences in ability. (In Appendix 5, there is a brief description of a small experiment carried out in the context of the interview to test whether or not stating the probability of success explicitly had any effect on the performance on the DSS forms. It did not.)

Thus the motivation measures included in the interview varied across several components of achievement (future orientation, power, autonomy, mastery, and so on) and across several assessment methods (story-telling, forces choice, and actual behavior).

Practical Considerations

The theoretical discussion of measurement above implies extensive analyses in terms of dividing the items into component measures, dividing the sample into homogeneous groups, and finding optimal combinations of measures for predicting various behaviors. However, these implications had to be balanced against the economic and time constraints upon both the interview and the data analyses. To reduce the number of possible analyses and make the most efficient use of limited data, the following assumptions and arbitrary decisions were made.

1. Projective achievement imagery and risk-taking behavior were assumed to be the best measures of achievement present in the interview. These two would be the criteria against which prospective survey measures would be tested. These were adopted for two major reasons. First, they do seem to be clear umbrella measures, as we pointed out earlier. As such, they should give considerable latitude in a search for valid new

techniques. Secondly, both are firmly entrenched in achievement motivation theory and research as proxies for the general motive. Even though both purport to measure this same entity, the methods differ so totally as to make selection of one over the other impossible. Thus the dual criterion might be expected to relate significantly to both of these criteria.

2. Population. All analyses would be run separately for the four race-sex groups: white males, black males, white females, and black females. This break-down was made for two reasons: first, race and sex seem by far the most important demographic factors affecting achievement, and second, these are distinctions which can be easily made on any survey sample, whereas others (IQ, SES, age) might not be.

3. To qualify as a "component measure", a scale must be composed of three or more "objective" items feasible for use in a national survey.

4. To qualify as a "component measure", a scale must either:

a. include items which a priori inspection shows to be related to one another and which meet some minimal criterion of internal consistency, or

b. include items empirically related to one another in the sub-sample and appear to make some theoretical sense.

These two methods of component construction are seen as equally important. The first, or "a priori" method utilizes past work on the components of achievement motivation. The second, or "empirical" method allows leeway for variation in the number and content of components across sub-samples.

5. Component measures would be constant within any one group. That is, a black male power measure would not change with the behavior to be predicted, but remain constant for all analyses.

6. To qualify for recommended use, a component measure must show a definite relation to either risk-taking or projective achievement imagery, preferably both.

7. Linear combinations of equally weighted components would be assumed.

Plan of Analysis

In keeping with the above assumptions, the data analysis was divided into three major steps:

1. Criteria development. For each criterion, several measures were available: items P1, P2, P3, and P4 for projectives; five post-intelligence test choices and Digit Symbol Substitution choice and differential performance for risk-taking. Correlational data and past research findings were used to select the best of these measures which were then formed into two criterion scales. By analysis of variance, the scales were checked for contaminants; effects of race of interviewer, IQ, or length of projective response; and interactions of these with the respondent's race and sex. Any contaminating effects were removed by regression.

2. Component development. Five *a priori* scales were formed from the 34 items by inspection. This process was not done separately for race-sex groups. The empirical scales were formed by means of a factor analysis of the 34 items for each group and selection of the "defining" items from each factor for each group. The internal consistency of the items on the scales was examined.

3. Testing the usefulness of component measure. Each component measure was then tested for its relation to the two criteria.

Selection of scales for use in a national survey was based on the analyses described above, plus some practical considerations about administrative difficulties and the results of an investigation of the scales' heuristic value in relating to various social status variables. These factors will be further discussed in the chapter on recommendations.

CHAPTER FIVE

METHODS OF ASSESSING ACHIEVEMENT MOTIVATION IN A SURVEY: ANALYSIS OF DATA

Criteria Development

Projective Criterion

Six projective items were administered; their intercorrelations, computed separately for **males and females**, are shown in Table 5.1. Items SP1 and SP2, the two incomplete sentences, were rejected from the criterion because of their negative correlations with the other items. Item P3 was also rejected, for two reasons: (1) its average correlations to items P2 and P4, although positive, were lower than the other average intercorrelations, and (2) this item was not coded on the standard 0-11 scale, but on a 0-2 abbreviated scale. Thus, the three items coded by the standard method (an 11-point scale developed by Atkinson, 1958), P1, P2, and P4, were summed to form the projective criterion scale. Because previous research (Veroff, et al., 1960) had found that the sheer length of response affects the amount of achievement imagery, an analysis of variance with response length as covariate was run. The results, in Table 5.2, show length of response, interviewer race, and sex to be significant effects, with no significant interactions: the respondents with black interviewers had more achievement imagery,* while women gave more achievement imagery than men. (Since different stimulus pictures were employed for the two sex groups, this difference was not of concern). However, the effects of interviewer race and response length (quantified by a regression equation, as shown in Table 5.3) had to be removed from the projective criterion. Therefore, a new score, the projective residual, was computed for each

---

*This was in spite of the fact that the black interviewers tend to elicit shorter stories in general.

Table 5.1

Intercorrelations of Projective Assessments of Achievement Motivation, by Sex

		Stories Told to Picture Cues		Stories Told to Verbal Cues		Sentence Completions	
		P(1)	P(2)	P(4)	P(3)	SP1	SP2
P(2)	Male	.02					
	Female	.28**					
P(4)	Male	.14*	.08				
	Female	.06	.11				
P(3)	Male	.15*	.09	.04			
	Female	.07	.02	.11			
SP1	Male	.13	.11	.07	-.06		
	Female	.14*	.00	-.08	.09		
SP2	Male	.01	-.07	-.01	.01	-.03	
	Female	.03	.11	-.11	.25**	.27**	

N Males, 173
N Females, 184

*$p \leq .05$
**$p \leq .01$

Table 5.2

Analysis of Covariance of Achievement Imagery on P1, P2, and P4, with Length of Response as Covariate: Race of Respondent x Race of Interviewer x Sex

Mean Achievement Imagery, Standardized + Adjusted for Length of Response

		Race of Respondent			
		White		Black	
		Male	Female	Male	Female
Race of Interviewer:	White	-.37	-.03	-.47	-.03
	Black	.16	.69	-.11	.34

Analysis of Covariance

	df	Sum of Squares	F
Interviewer Race	1	99.9	11.03**
Respondent Race	1	2.5	0.28
Respondent Sex	1	98.7	10.90**
Interviewer x Respondent Race	1	9.2	1.01
Interviewer Race x Sex	1	18.3	2.03
Sex x Respondent Race	1	1.9	0.21
Interviewer Race x Respondent Race x Sex	1	0.6	0.06
Regression (Words)	1	259.2	28.62**
Adjusted Error	3	3196.9	---

**$p \leq .01$

Table 5.3

Regression on Achievement Imagery on P1+P2+P4, Race of Interviewer and Length of Response as Predictors

Predictor	β
Race of Interviewer	.169
Length of Response	.262

Total variance explained: 9%

respondent by subtracting his predicted score based on the regression analysis in Table 5.3 from his raw score. This residual, then, is free of contaminating effects and is the Projective Criterion to be used in further analyses.

Risk-Taking Criterion

Seven different assessments of risk-taking were included in the interview: choices of easy, difficult, or moderate items after the Digit Span, Sentence Completion, Raven, Information, and Ammons tests; choice of an easy or moderate Digit Symbol Substitution form; and the difference in performance on DSS moderate and DSS long forms. The intercorrelations of these measures, by sex groups, are shown in Table 5.4. The Digit Span and DSS choices and DSS performance differential do not correlate among themselves or with the other four choices; however, the SCT, Raven, Ammons, and Information choices do form a cluster of significant correlations. Accordingly, these four items were selected for the scale.

The scale is the number of moderately-difficult choices made by the respondent in four opportunities. Choices were made in response to the following query from the interviewer:

"I've scored your performance so far on this test. Let's try just one more of these sentences. Which one would you like to try?

a. One that should be easy for you,
b. One that should be hard for you, or
c. Something in between?"

A "c" choice was met by another choice, between d. "One like most of the ones you got right," or e. "One a little harder than most of the ones you got right?"

The wording of these questions to emphasize a difficulty range tailored to the respondent's performance—the arrangement of test items from easy to hard and the administration of a further item actually contingent upon that performance level (i.e., a respondent who scored high on a test and then chose alternative b was given the most difficult item)-were all

Table 5.4

Intercorrelations of Behavioral Risk-taking Measures, by Sex

		Digit Span (1)	SCT (2)	Raven (3)	Ammons (4)	Information (5)	DSS (6)	Differential DSS Performance (7)
(2)	Male	-.03						
	Female	.03						
(3)	Male	.09	.22**					
	Female	.18*	.26**					
(4)	Male	.11	.41**	.40**				
	Female	.15	.17	.39**				
(5)	Male	.15	.29**	.44**	.55**			
	Female	.20	.24*	.23*	.37**			
(6)	Male	.03	.02	.05	.07	.15		
	Female	.04	-.06	.12	.10	.16		
(7)	Male	-.09	-.04	-.11	-.18	-.31**	-.01	
	Female	.01	.01	.02	.11	-.07	-.10	

$^{*}p \leq .05$
$^{**}p \leq .01$

intended to equalize expectations of success across respondents and tests. That is, it was hoped that all R's choosing a given alternative would consider the probability of success on that item the same. This probability should have between .80 and 1.00 for easy, .60 and .80 for moderately easy, .30 and .60 for moderately difficult, and .00 and .30 for difficult.

Unfortunately, the success expectations apparently were not equalized across all subjects but were related to test performance. Table 5.5 shows the percent of low and high IQ R's choosing the difficult alternative on all four opportunities; Table 5.6 shows the correlations between difficulty of choice and Sentence Completion test performance. Clearly, the high-IQ respondents did not believe that the choices were tailored to their own performance, and had a higher expectation of success than did the low-IQ respondents.

Since our assumption that there are equal expectancies across respondents is crucial if the risk-taking measure is to reflect achievement motivation, a correction for the contaminants like that for the projectives was followed. Table 5.7 shows the results of an analysis of variance on choice difficulty, with IQ as covariate. Again, interviewer race and respondent sex, in addition to IQ, affect the difficulty scale. The interviewer race effect is very strong: white interviewers elicited more difficult choices, regardless of race of respondent. This result is considered in full detail in Appendix 4.

The sex difference seems to be a real difference rather than a methodological difference. Much previous research has demonstrated that women are less likely to take extreme risks than men (see Veroff, 1969, and Crandall, 1969). Therefore, we did not view the sex difference in risk choice as calling for a correction. Furthermore, we are considering all analyses separately for males and females.

However, the IQ and race of interviewer effects do seem to call for correction in the risk-taking assessment. Since these effects were not ones that seem to reflect interactions with other variables, their direct effects on risk-taking were removed by the use of the regression equation.

Table 5.5

Percent of Respondents Choosing All "Very Difficult" Choices, by Sex and Intelligence Measures

	Low I.Q.	High I.Q.
Males	20%	37%
Females	13%	22%

Table 5.6

Correlation between I.Q. (Sentence Completion Test) and Mean Difficulty of Choice, by Sex and Race Groups

Group	r	N
Males	.28**	178
Females	.37**	187
Whites	.37**	177
Blacks	.24*	188

*p $\leq$ .05
**p $\leq$ .01

Table 5.7

Analysis of Covariance of Mean Difficulty in Risk Preference, with I.Q. as Covariate
Race of Respondent x Race of Interviewer x Sex

Mean Difficulty of Risk, Standardized and Adjusted for I.Q.

		Race of Respondent			
		White		Black	
		Male	Female	Male	Female
Race of Interviewer:	White	.28	.01	.96	-.15
	Black	-.28	-.59	-.22	-.30

Analysis of Covariance

	df	Sum of Squares	F
Interviewer Race	1	282.814	25.455**
Respondent Race	1	2.584	0.233
Respondent Sex	1	160,486	14.445**
Interviewer x Respondent Race	1	1.328	0.120
Interviewer Race x Sex	1	3.905	0.352
Sex x Respondent Race	1	0.478	0.043
Interviewer Race x Respondent Race x Sex	1	21.868	1.968
Regression (I.Q.)	1	557.615	50.188**
Adjusted Error	354	3933.104	---

**$p \leq .01$

Table 5.8 shows the regression equation used to predict difficulty of choice; the difference between predicted and raw score is residual risk difficulty.

Table 5.8

Regression on Mean Difficulty of Choice, Interviewer Race and IQ as Predictors

Predictor	β
IQ	.373
Race of Interviewer (Black Coded 1) (White Coded 0)	.243

Total Variance Explained: 18%

This new scale, now uncontaminated by the effects of I.Q. and interviewer race, should have a curvilinear relation to achievement motivation: respondents with high achievement motivation should, according to the theory, have moderate score on the scale, and respondents with low achievement motivation should have scores at the extremes. In order to use the scale more easily in further correlational analysis, it was "folded" by forming a new scale that gives high scores to moderate risk choices and low scores to either extremely low or high risk choices.

Table 5.9 shows how these folded moderate risk scores were assigned to each point on the distribution. To do this, we looked for that level of moderate risk-taking correction that would give us a mean and modal value in the uncorrected scores that was equivalent to choosing all moderate risk (8). This occurred for a category of -.25 to 2.00, and so we assigned this the highest value on the corrected scores and then gave progressively lower scores to levels folding from that point, as can be seen in Table 5.9. Those assigned "0" in the risk-taking criteria had a

Table 5.9

Relation of Risk-taking Criterion, Corrected Risk Difficulty, and Uncorrected Risk Choice Scales

Score on Risk-taking Criterion[1]	Score on Corrected Risk Difficulty[2]	Uncorrected Risk Choice[3] Range	Mean	Mode	N
0	-10.00 to -3.50	0-6	2.10	2	63
1	- 3.50 to -1.75	2-8	4.35	4	52
2	- 1.75 to -0.25	2-9	6.00	6	52
3	- 0.25 to 2.00	4-12	8.65	8	86
2	2.00 to 3.00	8-12	10.63	10	38
1	3.00 to 4.25	8-12	11.05	12	38
0	4.25 to 10.00	10-12	11.58	12	33

[1] Is a recoding of (2) in which score is linearly related to the "moderateness" of the risk choice.

[2] Is the difference between (1) and the regression equation prediction of (1), based on interviewer race and I.Q.

[3] Is the sum of scores on 4 separate choices (Sct, Raven, Ammons, Information), where each choice is coded 0 (easy), 1(moderately easy), 2 (moderately difficult), and 4 (difficult). Thus all easy would be scored 0; all moderate, 8; all hard, 12.

mean risk-taking score that was close to being all easy choices or all difficult choices. This new scale will become the Risk-Taking Criterion.

It is important to note in passing that the two criteria we have isolated show a significant positive correlation in males (+.19) but an insignificant and negative (-.04) correlation in females (both race groups in each case). Although this is further evidence for the poor generalizability of achievement motivation research to female populations, the positive correlation found in males was reassuring. Since it is so low, we felt that a combined criterion was not justified, and proceeded to use each criterion separately, expecting that a promising new technique should relate to one or the other criterion, and hopefully both.

Component Development

A Priori Component Measures

The ten paired-comparison, seven triad, three interview experience, and two semi-projective items were divided into four *a priori* scales: Mastery,(competition with a self-set standard of excellence; feelings of competence), Future Orientation (long-term goals), Autonomy-Power (ability to do things on one's own and to influence others), and Social Comparison (competition with others). These scales are displayed in Table 5.10. Also considered *a priori* scales were the Past Test Anxiety scale of seven items and the Personal Efficacy Scale, a five-item scale derived from the work of Epps, Gurin, and Morgan (Epps, 1969; Gurin, *et al.*, 1969; Morgan and Smith, 1970). These are shown in Tables 5.11 and 5.12.

To test internal consistency, alpha coefficients were computed for each scale, where

$$\alpha = \frac{\text{(average intercorrelation r) x no. of items n}}{1 + \text{ r x (n-1)}} .$$

These are shown in Table 5.13. The range of coefficients is adequate for all the scales except Future Orientation and Social Comparison, which display virtually no internal consistency.

Empirically-Derived Component Measures

A principal components factor analysis with varimax rotation was computed for each race-sex group; factors were extracted until 50% of the variance was "explained". The 34 items included, plus the abbreviated description of each used in listing the factors scales, are listed in Table 5.14.

For each factor, all items with loadings of .40 or more were taken as "defining"; in addition, items not loading over .40 on any factor were assigned to the factor on which they loaded highest. In forming the scales from these defining items, a few items which made no theoretical (or common)

Table 5.10

A *Priori* Achievement Motivation Scales

Item Symbol	Point Credit in Scale	Item Alternatives
MASTERY SCALE		
pc3		Why would you quit a job?
	0	It was too difficult after all
	2	It was not challenging any more
pc7		Which would you rather hear about yourself?
	0	Other people like him very much
	2	He can do anything he sets his mind on doing
pc8		Which would you rather hear about yourself?
	0	He can easily getother people to do things for him
	2	He is living proof that hard work gets a person ahead
pc9		Which would you rather hear about yourself?
	0	He is a really dependable and loyal friend
	2	He is always ready to meet challenges that come his way
ie2		If you learn you are doing well on the tests, would you
	0	Feel good about that
	2	Want to know more about the tests
tlc		Would you want your child to want very much to do his best
	0	Least
	1	Neither (alternatives: stand up for his rights, show kindness toward his playmates)
	2	Most
t5b		I would like to do better at what I try
	0	Least
	1	Neither (alternatives: have more friends, have other people pay attention to my point of view)
	2	Most
t6c		Would you like a job where you can show your real skills
	0	Least
	1	Neither (alternatives: where no one bosses you around, where you have lots of friends)
	2	Most
FUTURE ORIENTATION SCALE		
pc4		What kind of work do you like best?
	0	Start fresh every day
	2	Have to plan for the future and follow up on plans
ie3		If you learn you are doing well on the tests, would you
	0	Feel good about what you've done so far
	2	Mostly think about the tests yet to come
t3a		Would you like your child to learn to plan ahead for the future
	0	Least
	1	Neither (alternatives: not to appear weak to others; to do what his teachers expect of him)
	2	Most

continued

Table 5.10 -- Continued

Item Symbol	Point Credit in Scale	Item Alternatives
spl	Response to "In the next ten years I'm going to _____."	
	0	No achievement imagery
	1	Doubtful or low achievement imagery
	2	Definite achievement imagery
AUTONOMY-POWER SCALE		
pcl	Why would you quit a job?	
	0	Had too many things to decide about
	2	Had too little to say about how things got done
pc5	What do you think about after a hard job?	
	0	Whether the boss thought you did a good job
	2	whether you thought you did a good job
t7a	Would you want a job where you had to think for yourself?	
	0	Least
	1	Neither (alternatives: people you work with are nice group,
	2	Most you have a lot to say in what's going on)
t4b	Would you want your child to be a leader?	
	0	Least
	1	Neither (alternatives: to do the work hie teacher expects,
	2	Most to be popular with his classmates)
pc6	Which would you rather overhear about yourself?	
	0	People like to live next door to him
	2	His opinion carries a lot of weight among people who know him
pcl0	Which would you rather overhear about yourself?	
	0	He is fun to have at a party
	2	People like to go to him for advice on important matters
SOCIAL COMPARISON SCALE		
pc2	Why would you quit a job?	
	0	Too much competition among the workers
	2	You were a much better worker than anyone else
iel	What would you like to know about the tests	
	0	How well you did compared to others in the country
	2	How well you did compared to someone of your background
t2b	Would you like your child to do as well as most kids his age?	
	0	Least
	1	Neither (alternatives: make friends easily, defend himself
	2	Most if attacked)
sp2	Response to "compared to other men/women my age, I ______."	
	0	No achievement imagery
	1	Doubtful or low achievement imagery
	2	Definite achievement imagery

Table 5.11

Past Test Anxiety Scale

1 After you finished an important test, how confident did you feel about how you'd done?

0	1	2
not very confident	about average	very confident

2 When you took a final exam, did you have an uneasy, upset feeling?

0	1	2
was very upset	was somewhat upset	was not upset at all

3 When you were working on important tests, how fast did your heart beat?

0	1	2
beat very fast	faster than normal	about normal

4 How much do you think your emotions hurt your performance on the tests?

0	1	2
hurt it very much	hurt it slightly	did not hurt at all

5 When you were taking an important examination, how much did you perspire?

0	1	2
a great deal	more than usual	not at all

6 During tests, how much did you worry about what it would mean to fail?

0	1	2
worried a lot	worried some	did not worry at all

7 During a final examination, did you get so nervous that you couldn't remember some things you really knew?

0	1	2
forgot most of what I knew	forgot a few things I knew	did not forget anything I knew

Table 5.12

Personal Efficacy Scale

Item Symbol	Point Credit in Scale	Item Alternatives
PE.1	1	What happens to me is my own doing.
	0	Sometimes I feel that I don't have enough control over my life.
PE.2	0	Many times, I might just as well decide what to do by flipping a coin.
	1	In my case, getting what I want has little or nothing to do with luck.
PE.3	1	I've usually felt pretty sure my life would work out the way I want it to.
	0	There have been times when I haven't been very sure that my life would work out the way I want it to.
PE.4	1	I am the kind of person who plans his life ahead all the time.
	0	I am the kind of person who lives more from day to day.
PE.5	1	When I make plans ahead, I usually get to carry things out the way I expected.
	0	When I make plans ahead, things usually come up to make me change my plans.

Table 5.13

Coefficients of Internal Consistency (α)
for the *A Priori* Scales

Scale	White Males	Black Males	White Females	Black Females
Autonomy-Power	.22	.35	.40	.18
Social Comparison	.07	.04	-.25	.00
Future Orientation	.04	.04	.00	.20
Mastery	.24	.42	.25	.38
Personal Efficacy	.41	.29	.31	.19
Anxiety	.70	.67	.58	.76

Table 5.14

Items Used in Factor Analysis,
With Abbreviations

Symbol	Abbreviation (Item)
pc1	Too little say so quit job/too much to decide on Suppose you were on a job and you were thinking of quitting. Why do you suppose you'd quit? Which of these would come closest to describing you? a. You had too little to say about how things got done, or b. You had too many things to decide about.
pc2	Too much competition so quit job/better worker Which of these two would come closer to describing why you might quit a job? a. Too much competition among workers on the job, or b. You were a much better worker than anyone else?
pc3	Job too difficult so quit/job no challenge How about these? a. The job was too difficult after all, or b. The job was not challenging anymore?
pc4	Plan for future/start fresh every day Now which of these two would describe what kind of work you like best? a. Work where you have to plan for the future and follow up on plans, or b. Work where you can start on something fresh every day?
pc5	Boss thought you did a good job/you did If you'd finished a hard job, which would you think about? a. Whether your boss thought you did a good job, or b. Whether you thought you did a good job?

Table 5.14 -- Continued

Symbol	Abbreviation (Item)
pc6	Weighty opinion/good neighbor Which would you rather overhear about yourself? a. His/Her opinion carries a lot of weight among people who know him/her, or b. People like to live next door to him/her.
pc7	Other people like him/can do whatever he wants Now these two. Which would you rather hear about yourself? a. Other people like him/her very much, or b. He/She could do anything he/she sets his/her mind on doing.
pc8	Can get others to act/hard work gets ahead Now these two... a. He/She can easily get other people to do things for him/her, or b. He/She is living proof that hard work gets a person ahead.
pc9	Ready to meet challenges/good friend Now these two... a. He/She is always ready to meet challenges that come his/her way, or b. He/she is a really dependable and loyal friend.
pc10	Party fun/gives important advice Now these two... a. He/she is fun to have at a party, or b. People like to go to him/her for advice on important matters.
tlc	Want to do his best/stand up for own rights, be kind What would you like your child to do most? Which of the three would you least like him to do? a. Stand up for his own rights, b. Show kindness toward his playmates, c. Want very much to do his best.

Table 5.14 -- Continued

Symbol	Abbreviation (Item)
t2a	Make friends easily/do as well as others, defend self How about these? Which of these would you most like your child to do? Which of the three would you least like him to do? a. Make friends easily, b. Do as well as most kids his age, c. Defend himself if attacked.
t3c	Do what teachers expect/plan ahead, not be weak How about these? Which of these would you most like your child to learn? Which of the three would you least like him to learn? a. To plan ahead for the future, b. Not to appear weak to others, c. To do what his teachers expect of him.
t4b	Be a leader/do what teacher expects, be popular How about these? Which of these would you most and least like to teach your child? a. To do the work his teacher expects, b. To be a leader, c. To be popular with his classmates.
t5b	Do better at what I try/more friends, get attention Which is truest for you? Which of the three is least true? a. I would like to have more friends, b. I would like to do better at what I try, c. I would like to have more people pay attention to my point of view.
t6b	Lots of friends/no bosses, show real skills If you were to take a job -- maybe even something quite different from what you're doing now -- what would you most want? Which of these would you want least? a. A job where no one bosses you around too much, b. A job whereyou have a lot of friends, c. A job where you can show your real skills.

Table 5.14 -- Continued

Symbol	Abbreviation (Item)
t7b	Nice group of workers/think for self, lots of say on job How about these aspects of a job? Which of these would you want most? Which of the three would you want least? a. A job where you had to think for yourself? b. A job where the people you work with are a nice group? c. A job where you have a lot to say in what's going on.
ie1	Compare self to others in nation/of own background What would you like to know most about the tests? a. How well you've done compared to most others in this country, or b. Just how well you have done for someone of your background?
ie2	Want to know more about tests/feel good Suppose I told you that you were doing very well on the tests. a. Would you want to know more about the tests, or b. Would you feel good about that?
ie3	Feel good/think of tests to come Suppose I told you that up until now you were doing very well on these tests. a. Would you mostly feel good about what you've done so far, or b. Would you mostly think about the tests yet to come?
pe1	What happens is my own doing/not enough control over life a. What happens to me is my own doing. b. Sometimes I feel that I don't have enough control over my life.
pe2	Could decide by coin flip/no luck involved in life a. Many times, I might just as well decide what to do by flipping a coin. b. In my case, getting what I want has little or nothing to do with luck.

Table 5.14 -- Continued

Symbol	Abbreviation (Item)
pe3	Life will work out OK/not so sure a. I've usually felt pretty sure my life would work out the way I want it to. b. There have been times when I haven't been very sure that my life would work out the way I want it to.
pe4	Plan ahead/live day-to-day a. I am the kind of person who plans his life ahead all the time. b. I am the kind of person who lives more from day to day.
pe5	Plans work out OK/plans change a. When I make plans ahead, I usually get to carry things out the way I expected. b. When I make plans ahead, things usually come up to make me change my plans.
A1	Confidence After you finish an important test, how confident did you feel about how you'd done?
A2	Uneasy, upset feeling When you took a final exam, did you have an uneasy, upset feeling?
A3	Heart beat When you were working on important tests, how fast did your heart beat?
A4	Emotions How much do you think your emotions hurt your performance on the tests?
A5	Perspire When you were taking an important examination, how much did you perspire?

Table 5.14 -- Continued

Symbol	Abbreviation (Item)
A6	Worry about failure During tests, how much did you worry about what it would mean to fail?
A7	Nervous forgetting During a final examination, did you get so nervous that you couldn't remember some things you really knew?

sense were deleted. The resulting scales are shown in Tables 5.15 through 5.18. Although most of the scales are very short, the alpha coefficients are quite high, as would be expected for empirically-derived scales. For many of the scales there seem to be few distinguishing characteristics in the scales' contents; many are a mixture of the theoretical components used in forming the _a priori_ scales. Therefore, rather than labeling each factor, we will closely examine only those that correlate significantly with the criteria.

Testing the Scales

Table 5.19 reports the correlations of each _a priori_ and empirical scale with the projective and risk-taking criteria in each race-sex group.

A Priori Scales

Several of the _a priori_ component measures correlate highly with the criteria. Anxiety should correlate negatively, and does, especially for black males. The Mastery scale, while showing no extremely high correlations, relates consistently positively to both criteria, as does the Autonomy-Power scale. The Future Orientation measure shows little consistency across groups, although it too shows a general positive association reaching significance in white females. The Social Comparison scale seems useless except that it is the best predictor for risk-taking in white males, suggesting that white males who take such an orientation tend not to be moderate risk-takers. Finally the Personal Efficacy measure shows a very interesting pattern of correlations: positive to both criteria for white males, but _negative_ to projective imagery and positive to risk-taking in the other three groups. These results will be discussed more fully in a later section.

Empirical Scales

Although many of the empirical scales correlated with neither criterion, a few do appear promising. Let us briefly discuss the scales that correlate effectively with either or both criteria for each of the sub-groups separately.

Table 5.15

Empirical Scales Resulting from Factor Analyses, White Males

Scale, Alpha	Items	Achievement alternative/Other alternative	Factor Loadings
WM1	A2-A6	(All anxiety items)	49-73
.72			
WM2	pc8	Can get others to act/Hard worker	50
.48	ie1	Compare to others in country/Own background	50
	ie2	Want to know more/Feel good	66
	sp1	High achievement-next 10 years/Low	55
WM3	pc4	Job to plan and follow up/Start fresh	71
.52	pe1	What happens is own doing/No control	71
	pe2	No luck involved in life/Coin flip	34
WM4	t1	Wants to do best/Own rights, show kindness	76
.32	pe2	No luck involved in life/Coin flip	36
	sp2	High achievement-compared to others/Low	60
WM5	pc2	Better worker so quit/Too much competition	48
.51	pc3	Job no challenge so quit/Too hard	56
	t7	Job to think for self/Nice group, lots of say	56
WM6	t2	Do as well as others/Make friends, defend self	56
.42	53	Plan ahead/Not be weak, teacher expects	49
	ie3	Think of tests to come/feel good	46
	pe4	Plan ahead/Live day-to-day	63
WM7	pc6	Weighty opinion/Good neighbor	62
.39	t4	Leader/Teacher expects, be popular	59
	pe5	Plans work out/Plans change	57
WM8	pc1	Too little say so quit/Too many decisions	60
.48	pc5	You thought you did a good job/Boss did	37
	t6	Show skills/No boss, lots of friends	76
WM9	pc7	Can do whatever he wants/Others like him	72
.49	pc9	Challenge meeter/Good friend	60
	pe3	Life will work out OK/Not sure	48

Table 5.16

Empirical Scales Resulting from Factor Analyses, Black Males

Scale, Alpha	Items	Achievement alternative/Other alternative	Factor Loadings
BM1	A1	Confidence	.71
.60	A2	Upset feeling	.42
	A4	Emotions	.60
BM2	pc1	Too little say so quit/Too many decisions	.38
.51	pc5	You thought you did a good job/Boss did	.71
	pc7	Can do whatever he wants/Others like him	.69
	t4	Leader/Teacher expects be popular	.50
BM3	A2	Anxiety upset feeling	.41
.67	A3	Heart beat	.42
	A5	Perspire	.69
	A6	Failure worry	.46
BM4	pc6	Weighty opinion/Good neighbor	38
.38	pc10	Gives important advice/Party fun	41
	ie2	Wants to know more/Feel good	64
	ie3	Think of tests to come/Feel good	62
BM5	pc4	Job to plan and follow up/Start fresh	45
.50	pc8	Hard worker/Can get others to act	64
	pe2	No luck involved/Coin flip	60
	pe4	Plan ahead/Live day-to-day	64
BM6	pc9	Challenge meeter/Good friend	64
.36	T7	Thinks for self/Nice group, lots of say	49
	sp2	High achievement-compared to others/Low	41
	pe5	Plans work out/Plans change	40
BM7	t5	Do better/More friends, pay attention	61
.47	t1	Wants to do best/Own rights, show kindness	62
	pe1	What happens is own doing/No control	47
BM8	pc2	Better worker so quit/Too much competition	76
.57	pc3	Job no challenge so quit/Too hard	62
	t5	Do better/More friends, pay attention	36
	t6	Show skills/No boss, lots of friends	61

Table 5.17

Empirical Scales Resulting from Factor Analyses, White Females

Scale, Alpha	Items	Achievement alternative/Other Alternative	Factor Loadings
WF1 .58	A1-A7 (All anxiety items)		
WF2	pc5	You thought you did a good job/Boss did	40
.44	pc7	Can do whatever he wants/Others like him	57
	pc9	Challenge meeter/Good friend	49
	t7	Think for self/Nice group, lots of say	67
WF3	pc5	You thought you did a good job/Boss did	45
.56	pc6	Weighty opinion/Good neighbor	35
	t4	Leader/Teacher expects, be popular	73
	ie2	Want to know more/Feel good	44
	pe5	Plans work out/Plans change	55
WF4	pc8	Hard worker/Can get others to act	63
.54	t5	Do better/More friends, pay attention	78
	pe1	What happens is own doing/No control	55
	pe2	No luck involved in life/Coin flip	41
WF5	t1	Wants to do best/Own rights, kindness	54
.30	t3	Plan ahead/Not be weak, teacher expects	34
	pe3	Not sure life will work out/Sure	43
	pe4	Live day-to-day/Plan ahead	66
WF6	pc2	Better worker so quit/Too much competition	52
.52	pc3	Job no challenge/Too hard	65
	pc10	Gives important advice/Party fun	65
WF7	sp1	High achievement-next 10 years/Low	67
.23	sp2	High achievement-compared to others/Low	47

Table 5.18

Empirical Scales Resulting from Factor Analyses, Black Females

Scale, Alpha	Items	Achievement Alternative/Other Alternative	Factor Loadings
BF1	A1-A7	(All anxiety items)	
.76			
BF2	pc4	Job to plan and follow up/Start fresh	50
.22	pc9	Challenge meeter/Good friend	41
	t3	Plan ahead/Not be weak, teacher expects	50
	pe2	No luck involved in life/Coin flip	46
BF3	pc7	Can do whatever he wants/Others like him	64
.60	pc8	Hard worker/Can get others to act	64
	t5	Do better/More friends, pay attention	64
	t7	Think for self/Nice group, lots of say	56
BF4	pc5	Boss though you did a good job/You did	57
.51	ie1	Compare to own background/to others in country	71
	sp1	High achievement- next 10 years/Low	54
	sp2	High achievement-compared to others/Low	52
BF5	pc10	Gives important advice/Party fun	68
.52	t6	Show skills/No boss, lots of friends	52
	ie3	Think of tests to come/Feel good	75
	pe5	Plans work out/Plans change	41
BF6	pc3	Job no challenge so quit/Too hard	52
.42	t2	Do as well as others/Make friends, defend self	58
	pe1	What happens is own doing/No control	67
BF7	pc2	Better worker so quit/Too much competition	53
.30	t4	Leader/Teacher expects, be popular	62
	pc6	Weighty opinion/Good neighbor	37

Table 5.19

Correlations of Empirical and *A Priori* Scales with Risk-Taking and Projective Criteria, by Race-Sex Groups

Group	White		Black		White		Black	
	Males		Males		Females		Females	
Criterion	Proj.	Risk	Proj.	Risk	Proj.	Risk	Proj.	Risk
Empirical Scales								
1	-.104	.002	-.091	-.049	-.214*	.045	-.080	-.059
2	.081	.218	.085	.008	.114	.375**	.135	.046
3	.262*	.017	-.242*	-.249*	.120	.244*	.007	.072
4	-.109	-.087	.126	.317**	.029	.038	-.006	.010
5	.170	.114	-.247*	-.004	.051	-.031	.215*	.197
6	-.030	-.026	-.009	.052	-.014	.151	-.227*	.023
7	.230*	.085	.129	.187	.049	.195	.106	.037
8	.026	-.039	.085	.044	----	---	---	---
9	.307**	.073	---	---	---	---	---	---
A Priori Scales								
Autonomy-Power	.133	.116	.118	.090	.057	.300**	.171	.102
Social Comparison	-.033	-.179	-.039	-.036	.001	-.018	-.124	.068
Future Orientation	.158	.100	.094	-.165	-.062	.318**	.142	.114
Mastery	.218*	-.012	.103	.159	.171	.147	.152	.073
Personal Efficacy	.328**	.054	-.184	.152	-.194	.103	-.179	.193
Anxiety	-.111	-.024	-.216*	-.255*	-.214*	.045	-.080	-.059

*$p \leq .05$

**$p \leq .01$

White males. Referring to Table 5.19, we can see that while no scale correlates significantly with the risk criterion for white males, three scales correlate significantly with the projective criterion (WM 3, WM 7, and WM 9). For convenience, the items in these factors with loadings over .40 are reproduced in Table 5.20 with our suggested identification names for the factors. For white males the empirical scales that seem to be effective (Planning, Power, and Mastery) overlap the *a priori* scales that are effective (Personal Efficacy, Autonomy-Power, and Mastery). Thus no new scales are suggested by the factor analysis for this group.

Black males. Two factors relate significantly to the projective criterion for the black males (BM 3 and BM 5), and two to the risk criterion (BM 3 again, and BM 4). Reproduced in Table 5.21 are the items with loadings of .40 or higher on these scales. BM 3 represents a selection from the test anxiety measure, which we have already seen is an effective predictor of the criteria in black males. The fact that only certain items are effective for the black males suggests that we should consider shortening the anxiety scale to those items for everyone. What is new from the current analysis is that BM 5 and BM 4 correlate with the criteria in an interesting pattern. Our interpretation of factor 5 is that it represents the achievement moral system of the dominant American society, with the high loadings on "hard work" coupled with "planning" and "rejection of fate control." Had the item identifying "hard work" (as opposed to "can get others to act") not been so high on this factor, we would have interpreted it as a straight efficacy factor. But that item's loading the way it does for these men suggests that the factor represents how much they have identified with the achievement value system as a moral system. And indeed for these men such an identification is negatively correlated with the achievement motivation criterion found in the projective instrument. It is as if black men who adopt these values as a moral system undermine their own motivations. A black male's agreeing with these items might be tantamount to agreeing with the cultural stereotype that the black man has only himself to blame if he does not achieve. Not agreeing with these moral statements might be

Table 5.20

Items with Loadings of .40 or Over on Empirical Scales Effective for White Males

Factor Number and Name	Scale Correlation to Projective Criterion	Scale Correlation to Risk-Taking Criterion		Items Positive Alternative/Other Alternative
WM3 Planning	.26	.02	PC4	Job to plan and follow up/start fresh
			PE1	What happens is own doing/no control
			PE2	No luck involved in life/coin flip
WM7 Power	.23	.09	PC6	Weighty opinion/good neighbor
			T4	Leader/teacher expects, be popular
			PE5	Plans work out/plans change
WM9 Mastery	.31	.07	PC7	Can do whatever he wants/others like him
			PC9	Challenge meeter/good friend
			PE3	Life will work out OK/not sure

Table 5.21

Items with Loadings of .40 or Over on Empirical Scales Effective for Black Males

Factor Number and Name	Scale Correlation to			Items
	Projective Criterion	Risk-Taking Criterion		Positive Alternative/Other Alternative
BM3 Anxiety	-.24	-.25	A2	Upset feeling during tests/not upset
			A3	Fast heartbeat during tests/not upset
			A5	Heavy perspiration during tests/no perspiration
			A6	Worried about failing tests/no worry
BM4 Instrumenal Orientation	.13	.32	PC6	Weighty opinion/good neighbor
			PC10	Gives important advice/party fun
			IE2	Wants to know more/feels good
			IE3	Think of tests to come/feel good
BM5 Moral Achievement	-.25	.00	PC4	Job to plan and follow up/start fresh
			PC8	Hard worker/can get others to act
			PE2	No luck involved/coin flip
			PE4	Plan ahead/live day-to-day

more like recognizing that the system might be at fault in accounting for the failure of black men. Indeed we think this scale indirectly measures how unwilling the person is to blame the American social system for his own lack of achievement. The fact that some items from the a priori Personal Efficacy scale appear on this scale for black males may account for the Personal Efficacy scale's tendency to relate negatively to the projective criterion in this group.

The positive effect of Instrumental Orientation (BM 4) in predicting the risk criterion is also provocative. We label it an Instrumental Orientation because in each of the items the positive item is a rejection of an affective, spontaneous type of response. The positive item thus seems to be a rejection of an immediate orientation to life and an endorsement of a more instrumental, manipulative orientation, sometimes stated in power terms (pc6 and pc10), sometimes as a future orientation (ie3), and sometimes in mastery terms (ie2). This suggests that the slight positive effectiveness of the a priori scales of Future Orientation, Mastery, and Autonomy-Power for black males rests in the extent to which these scales represent a rejection of the impulsive, the affective, and the more immediate orientation to life. Such an explanation will also be evident in the interpretation of some of the findings for black women.

White females. The scales resulting from the factor analysis that are effective in predicting either of the two criteria for white women are reproduced and identified in Table 5.22. It is quite clear from these results that we are getting a complete overlap between our a priori analysis for white females and this empirical analysis. Anxiety, Mastery, and Autonomy-Power are a priori scales similar to the factor scales identified in Table 5.11: Anxiety, Autonomy-Mastery, and Power (WF 1, WF 2, and WF 3). Thus for white females we learn little additional information from the factor analysis over our a priori analysis.

Table 5.22

Items with Loadings of .40 or Over on Empirical Scales Effective for White Females

Factor Number and Name	Scale Correlation to Projective Criterion	Risk-Taking Criterion		Items Positive Alternative/Other Alternative
WF1 Anxiety	-.21	.05		All seven anxiety items
WF2 Autonomy-Mastery	.11	.33	PC5	You thought you did a good job/boss did
			PC7	Can do whatever he wants/other like him
			PC9	Challenge meeter/good friend
			T7	Think for self/nice group
WF3 Power	.12	.24	PC5	You thought you did a good job/boss did
			PC6	Weighty opinion/good neighbor
			T4	Leader/teacher expects, be popular
			IE2	Want to know more/feel good
			PE5	Plans work out/plans change

Black females. Table 5.23 presents the two scales that are significant predictors of the criteria for the black females. We have called them Instrumental Orientation (BF 5) and Moral Achievement Orientation (BF 6), parallel to similar factors in the black males. The first, instrumental orientation, seems to be a composite of power, future orientation, and mastery, but the one thing in common across all items is an instrumental orientation: a rejection of an alternative suggesting a more affective, spontaneous, or immediate reaction to life. The Instrumental Orientation scale also correlated highest to the risk-taking criterion (.19), paralleling the results for the black males. The second scale that correlates significantly with the projective criterion for black females is what we have again called a Moral Achievement Orientation. Again, we would have looked at the scale as a measure of personal efficacy if we had just examined the item with the highest loading (pel: "what happens is my own doing" vs. "have no control over life"). Looking closer at the other items, however, we realized that in endorsing them a black woman rejects quitting a job because it is "too hard" and rejects wanting a child to defend himself if attacked. Considering the endorsement of achievement items in that context, we reappraised the interpretation of the measure as personal efficacy and considered it a measure of the adoption of the dominant cultural attitudes toward achievement in a moral sense. This attitude does not permit the black woman to see social inequities (a job too hard for her, or a world that might attack her children) as stemming from the system, but only as reflecting her own achievement inadequacies. As with the black males, we would see the adoption of such a code of achievement in blacks as a hindrance to personal achievement orientation, perhaps an obstacle standing in the way of their own efficacy, as the Gurins have suggested (1969).

Empirical and A Priori Scales Compared: A Summary

The factor analyses in the different groups led to some scales that were effective in predicting the criteria. Generally speaking, they overlapped with *a priori* scales we had established, especially in the white groups. However, the empirical scales that were effective in the black

Table 5.23

Items with Loadings of .40 or Over on Empirical Scales Effective for Black Females

Factor Number and Name	Scale Correlation to			Items
	Projective Criterion	Risk-Taking Criterion		Positive Alternative/Other Alternative
5 Instrumental Orientation	.22	.20	PC10	Gives importance advice/party fun
			T6	Show skills/no boss, lots of friends
			IE3	Think of tests to come/feel good
			PE5	Plans work out/plans change
6 Moral Achievement	-.23	.02	PC3	Job no challenge so quit/too hard
			T2	Do as well as others/make friends, defend self
			PE1	What happens is own doing/no control

groups lead to some special insights about the problems we were encountering in our a priori scales for these groups. Only certain items from the Anxiety Scale were effective for black males, suggesting that we use only those items for everyone. More importantly, Autonomy-Power, Mastery, and Future Orientation seem to be working in the black groups to the extent that they represent alternatives to the immediate, affective orientation to life (what we have called Instrumental Orientation), but not working to the extent that they represent moralistic values that would remove the blame for black failure from the social system to the individual (Moralistic Achievement Orientation). These conclusions suggest that the failure of the Personal Efficacy scale to be effective, and in some instances, actually relating negatively to the criteria of achievement orientation rests in the degree to which that scale can be viewed as a moralistic achievement orientation in certain groups. This was especially true for blacks, but there was some indication that such was the case for white females also. The fact that different items from the Personal Efficacy scale might have different moralistic connotations depending on a person's race or sex or both would suggest that it is not a generally viable scale for all groups.

These conclusions suggest that we stick with the a priori scales but "purify" them: first, shorten the Anxiety scale, then examine further the items in the three a priori scales that generally worked -- Autonomy-Power, Future Orientation, and Mastery -- to see if we can pull out items that avoid the moralistic orientation and perhaps accentuate the instrumental orientation.

Increasing the Effectiveness of the A Priori Scales

A new test anxiety scale. Since only certain items from the Test Anxiety scale were effective for the black males, we used only these items for everyone to see whether we could increase the effectiveness for black males and shorten the scale without significantly decreasing the effectiveness of the scale in other groups. The alpha coefficients and correlations of this Revised Anxiety Scale to the two criteria are presented for each group in Table 5.24. Neither the alphas nor the correlations are reduced significantly.

Table 5.24

Revised Anxiety Scale with Alpha Coefficients and Correlations to Criteria in Four Race-Sex Groups

The Revised Anxiety scale includes:

A2: When you took a final exam, did you have an uneasy, upset feeling?
A3: When you were working on important tests, how fast did your heart beat?
A5: When you were taking an important examination, how much did you perspire?
A6: During tests, how much did you worry about what it would mean to fail?

Group	N	Alpha Coefficient	Correlation to Projective Criteria	Correlation to Risk-Taking Criteria
White Males	(86)	.71	-.10	-.02
Black Males	(92)	.67	-.24*	-.23*
White Females	(91)	.71	-.14	.09
Black Females	(96)	.61	-.09	-.02

*p ≤ .05.

New scale for achievement orientation. In Table 5.25 we present the correlations between each of the items found on the three a priori scales with the two criteria. From that set of items we can reject those items that show any of the negative trends that may be due to the moralistic orientation we suggested in the previous section, and accept only those items that show a positive trend across most of the groups. We established the following criterion for including an item from one of these scales for a final scale: among the eight correlations of an item to the two criteria in the four groups, an item had to have a correlation greater than .10 in at least two of the groups on at least one criterion, and had to have no negative correlation less than -.10 in any group. With such a criterion four items from the Mastery scale, two from the Future Orientation scale, and four from the Autonomy-Power scale emerged as acceptable. These items will henceforth be grouped together and labeled as the Achievement Orientation scale; they appear in Table 5.26.

The internal consistency of the Achievement Orientation scale varied from .33 to .49 in the four groups. The coefficients are reported in Table 5.27, along with the correlations of the scale with the criteria in each group. We should note that the scale shows generally positive correlations in each group with each criterion, but is not particularly effective for black females.

A Resultant Achievement Orientation Scale

Giving equal weight to the new Achievement Orientation scale and Revised Anxiety scale, we subtracted the latter from the former, and called this scale the Resultant Achievement Orientation scale. This scale, its alpha for the four sex-race groups, and its effectiveness in predicting the two criteria are presented in Table 5.28. This is the most effective scale we have found across all groups--- again being the least effective for black females, but most effective for black males.

Since this is the scale to be recommended for general use in a heterogeneous sample, we thought it critical to see how this scale relates to other social characteristics of populations that might bear on issues of achievement orientation. In Table 5.29, we report for each of the sex-race

Table 5.25

Correlations of Item-to-Criteria for Items In Three *A Priori* Scales, by Race-Sex Group

Scale Item	White Males (N=86) Risk	White Males (N=86) Proj.	Black Males (N=92) Risk	Black Males (N=92) Proj.	White Females (N=91) Risk	White Females (N=91) Proj.	Black Females (N=96) Risk	Black Females (N=96) Proj.
Mastery								
PC3[1]	.13	.18	-.07	.03	.20	-.08	.02	.12
PC7[1]	.14	.23*	-.05	.15	.24*	.18	.00	.16
PC8	-.18	-.01	.14	-.18	-.17	.08	.08	.05
PC9	-.04	.21*	.01	-.11	-.12	.06	.02	.08
IE2[1]	.05	.00	.27*	.14	.24*	.17	.00	.01
T1c	.05	-.03	.07	.12	-.02	.03	-.03	-.04
T5b[1]	-.02	.06	.08	.12	.11	.07	.05	-.03
T6c	-.13	.03	-.04	.11	.14	-.07	.11	.23
Future Orientation								
PC4	.09	.21*	-.27*	.01	.21*	.01	.05	-.09
IE3[1]	.03	.05	.02	.04	.15	-.09	.12	.23*
T3	.04	-.02	-.11	-.18	.03	-.13	.04	.29**
SP1[1]	.05	.04	-.02	.22*	.17	.06	.00	-.05
Autonomy-Power								
PC1	.10	-.01	-.03	.13	.01	-.09	-.14	.20
PC5	-.02	.03	-.13	.02	.37**	.13	.19	.08
T7a[1]	.02	.08	.03	.15	.33**	-.06	.07	-.09
T4b[1]	.04	.15	.04	-.02	.14	.03	.10	.13
PC6[1]	.19	.03	.23*	.12	.02	.16	-.09	.12
PC10[1]	.06	.14	.19	-.05	.04	.03	.18	-.02

[1] Items meeting selection criterion

*p ≤ .05.

**p ≤ .01.

Table 5.26

Achievement Orientation Scale

Item	Positive Alternative/Other Alternative
PC3	Job no challenge so quit/job too difficult
PC6	Opinions hold lots of weight/people like him as neighbor
PC7	Can do whatever he wants/others like him
PC10	Gives important advice/fun at a party
T4	Be a leader/do what teacher expects, be popular
T5	Want to do better/more friends, get attention
T7	Think for self/nice group, lots of say
IE2	Want to know more about tests/feel good
IE3	Think of tests to come/feel good
SP1	High achievement response--next ten years/low

Table 5.27

Achievement Orientation Scale: Alpha Coefficients and Correlations to Criteria in Four Race-Sex Groups

Group	N	Alpha Coefficient	Correlation to Projective Criteria	Correlation to Risk-Taking Criteria
White Males	(86)	.33	.16	.16
Black Males	(92)	.42	.13	.16
White Females	(91)	.49	.10	.34**
Black Females	(96)	.49	.10	.05

**p ≤ .01.

Table 5.28

Resultant Achievement Orientation Scale:
Means, Sigmas, Alpha Coefficients, and Correlations
to Criteria in Four Race-Sex Groups

					Correlation to	
Group	N	Mean	Sigma	Alpha Coefficient	Projective Criteria	Risk-taking Criteria
White Males	(86)	16.7	3.3	.48	.18	.14
Black Males	(92)	16.9	4.1	.54	.23*	.26*
White Females	(91)	15.8	3.4	.50	.17	.25*
Black Females	(96)	15.5	3.4	.54	.12	.05

*$p \leq .05$

Table 5.29

Correlations of Resultant Achievement Orientation
to Selected Characteristics of Respondents
(by Sex-Race Group)

Social Characteristic	White Males (N=86)	Black Males (N=92)	White Females (N=91)	Black Females (N=96)
Education	.21*	.38**	.40**	.14
Income	-.06	.11	.15	.03
Occupational Mobility	-.05	.11	-.02	.05
Intelligence (SCT)	.27**	.37**	.32**	.12

* $p \leq .01$

** $p \leq .05$

groups the results of the correlations of the Resultant Achievement Orientation scale to education; income (reported family income by the respondent); occupational mobility (the difference between Duncan's status coding of the respondent's occupation and his father's;*) and our own best estimate of what is standardly measured as intelligence (the scores on the Sentence Completion Test). Quite apparent in that table is the consistent relationship of the scale to both education and the Sentence Completion Test, perhaps as it theoretically should be. The gnawing concern for us in thinking about these correlations is the possibility that the two criteria on which the scale depends -- the projective and risk-taking criteria -- are largely a function of these two other variables, education and intelligence, and nothing more than that. Remembering that the intelligence factor was systematically removed from the risk-taking criterion, we were really only concerned with the possible contamination of the education variable in accounting for the effectiveness of the Resultant Achievement Orientation scale. Our concern was mostly limited to the male groups, because only in them was there any substantial correlation between education and the criteria, as can be seen in Table 5.30 below. Therefore we recomputed the correlations between the Resultant Achievement Orientation scale and the two criteria in the four groups, partialling out the effects of education. These results are reported in Table 5.31 below.

Comparing Table 5.28 with Table 5.31, we can see that the effect of partialling out education in computing the correlation of the scale to the two criteria was negligible, with the largest reduction in the prediction of the projective criterion. We had some further evidence that helped remove our concern about the possible contaminating influences of education on our results for males. We checked out the correlations of the Resultant Motivation scale to the criteria in the males who were less educated (less than a high school degree), since this is a group that might be of greatest

---

*Women who did not work outside the home were not included; if the father's occupation was uncodable because the father was absent from the home, it was arbitrarily assigned to the next-to-lowest status coding.

Table 5.30

Correlation of Education to Projective and Risk-Taking Criteria (by Sex-Race Groups)

Group	N	Projective Criterion	Risk-Taking Criterion
White Males	(85)	.14	.32**
Black Males	(90)	.19	.26*
White Females	(90)	-.04	.11
Black Females	(94)	.15	.04

*p ≤ .05.

**p ≤ .01.

Table 5.31

Correlations of Resultant Achievement Orientation (A) to the Projective and Risk-Taking Criteria (C), Partialling out the Effects of Education (E) in Different Sex-Race Groupings

Group	N	$r_{AC \cdot E}$ Projective Criterion	$r_{AC \cdot E}$ Risk-Taking Criterion
White Males	(86)	.13	.12
Black Males	(92)	.14	.20
White Females	(91)	.20	.23*
Black Females	(96)	.11	.03

*p ≤ .05.

interest in pragmatic research and found that for this group the correlations to criteria were about as high as they were for the whole group; for white males, .24 with the projective criterion, and .13 with the risk-taking criterion; for black males, .16 for the projective criterion, and .23 with the risk-taking criterion. Thus, even in a group with a restricted range of education we find that the Resultant Achievement Orientation scale is effective.

As a final appraisal of the Resultant Achievement Orientation scale, we present in Table 5.32 the results of using this scale in combination with the intelligence measure recommended for general use in Chapter 3 to predict our major criterion of actual achievement in the society, education. As additional predictors we included status background of the respondent and age, since these factors are independently correlated with education in all race-sex groups considered. Table 5.32 presents therefore the effectiveness of our recommended motivational and intelligence measures, independent of each other and of the critical demographic variables. The results are very encouraging about the usefulness of the measure of resultant motivation in two groups -- white females and black males. We were already skeptical about its effectiveness for black females; the fact that it was also not effective for white males is very provocative, since it was on that group that most measures of achievement motivation were originally constructed. It might be that for white males the level of education attained is as much a function of "performing as expected", and responding to cultural norms as it is a function of individual achievement orientation. For the other groups -- those less advantaged in terms of opportunity for educational and occupational success -- individual achievement orientation does seem to be a significant factor in the prediction of an important performance variable, education.

Table 5.32

Regression on Education with Intelligence (SCT) and Resultant Achievement Orientation as Predictors along with Status Background and Age (By Sex x Race Groups)

	Group											
	White Males (N=86)			White Females (N=91)			Black Males (N=92)			Black Females (N=96)		
Predictor	Beta	R^2	t	Beta	R^2	t	Beta	R^2	t	Beta	R^2	t
Intelligence (SCT)	.45	.22	4.77**	.39	.17	4.17**	.48	.26	5.55**	.54	.33	6.64**
Resultant Achievement Orientation	.03	.00	0.32	.21	.06	2.25	.21	.07	2.50*	.07	.01	0.83
Status Background	.21	.06	2.24*	.15	.03	1.65	.18	.05	2.20*	.18	.05	2.25*
Age	-.09	.01	0.98	-.19	.05	2.11*	-.20	.07	2.07*	-.19	.06	2.33*

*p ≤ .05.

**p ≤ .01.

Summary

In our efforts to evolve a feasible, reliable, valid scale of motivation for achievement that could be applicable to a heterogeneous sample, we followed the following steps:

1. We developed two criteria for achievement orientation that have been standardly used in previous research and thinking about achievement -- score of achievement motivation from projective fantasy and the degree to which the respondent selected a moderate risk in a test of his skills. This process required some decontamination of confounding factors in assessment of the criteria.

2. We developed a set of a priori scales of achievement orientation from questionnaire material, based on previous research and thinking about achievement motivation.

3. We developed a set of empirical scales of these same questionnaire items based on their clustering in factor analyses. We did these analyses separately for sex-race groupings (white males, black males, white females, black females).

4. We tested the predictive effectiveness of each set of scales separately for the four different sex-race groupings.

5. We then "purified" the a priori scales for the four different sex-race groupings. While the results of testing the effectiveness of the empirical scales generally underscored the results obtained from the a priori analysis, the empirical scales did provide some insights into why the a priori scales might be in some error. In particular, the empirical results suggested that the a priori scale of Anxiety might be shortened to four items. Furthermore, the a priori scales of Power-Autonomy, Mastery, and Future Orientation might be effective to the degree that they account for an instrumental orientation to life in general as opposed to a more affective spontaneous orientation, and to the degree that they avoid a moralistic orientation to achievement, one that seems to put the blame for failure on the person rather than on the social system. To purify the a priori scales of Power-Autonomy, Mastery, and Future Orientation, we examined how each item within each of these a priori scales related to the

criteria, and then selected those items from the scales that showed consistent positive trends across the sex-ract groupings while showing no large negative trends. The ten items meeting our new criteria were combined into a revised Achievement Orientation scale.

6. The four items emerging from purifying the Anxiety scale in step 5, and the ten items emerging for purifying the other _a priori_ scales in step 5, were combined to form a Resultant Achievement Motivation scale. In this scale, the Revised Anxiety scale, equally weighted with the Revised Achievement Orientation scale, is subtracted from the latter.

7. The Resultant Achievement Orientation scale was found to correlate especially well with the criteria in black males, but not very effectively in black females. The results seem not to be dependent on the fact that both the scales and the criteria are correlated significantly with education.

8. Finally, we used the Resultant Achievement Orientation scale along with the intelligence measure most effective in a population (the Sentence Completion Test), status background, and age in a multiple prediction analysis of education. In these analyses the Resultant Achievement Orientation scale was found to be a significant independent predictor for black males and white females but not for black females or white males.

From these steps we feel we have a generally effective scale usable in a heterogeneous population; some misgivings about its usefulness with black females remain. Since we have some empirical scales found effective in black females, it may be that researchers interested in that group might want to consider these scales in addition to the Resultant Achievement Orientation scale.

REFERENCES

Angelini, A. L. Um novo metodo para avaliar a motivacao humano (A new method of evaluating human motivation). *Bol. Fac. Filos. Cienc. S. Paulo.*, 1955, No. 207. Cited in Lesser, G. D., Krawitz, R., and Packard, R. Experimental arousal of achievement motivation in adolescent girls. *Journal of Abnormal and Social Psychology*, *66*, 1963, 59-66.

Atkinson, J. W. Motivational determinants of risk-taking behavior. *Psychological Review*, *64*, 1957, 359-372.

Atkinson, J. W. A scoring manual for the achievement motive. In Atkinson, J. W. (ed.). *Motives in Fantasy, Action and Society*. Princeton: Van Nostrand, 1958.

Atkinson, J. W. *An Introduction to Motivation*. Princeton: Van Nostrand, 1964.

Atkinson, J. W., and Feather, N. T. (eds.) *A Theory of Achievement Motivation*. New York: Wiley, 1966.

Atkinson, J. W., and Moulton, R. W. *New Methods of Measuring Achievement Related Motives*. National Science Foundation Final Report. Grant #GS1399, 1969.

Birch, D., and Veroff, J. *Motivation: A Study of Action*. Belmont, California: Wadsworth, 1966.

Birney, R. C., Burdick, H., and Teevan, R. C. *Fear of Failure*. Princeton: Van Nostrand-Reinhold, 1969.

Dornbusch, S. M. Afterword. In Maccoby, E. E. (ed.). *The Development of Sex Differences*. Stanford: Stanford University Press, 1966.

Crandall, V. C. Sex differences in expectancy of intellectual and academic reinforcement. In Smith, C. D. (ed.). *Achievement-Related Motives in Children*. New York: Russell Sage, 1969.

Duncan, O. D. A socioeconomic index for all occupations. In Reiss, A. J., Jr. (ed.). *Occupations and Social Status*. New York: Free Press of Glencoe, 1961.

Epps, E. G. *Family and Achievement: A Study of the Relation of Family Background to Achievement Orientation and Performance among Negro High School Students*. Final Report, U. S. Department of Health, Education, and Welfare, Office of Education. Project #5-1006, 1969.

REFERENCES (Continued)

Fontana, G. J. An Investigation into the Dynamics of Achievement Motivation in Women. Unpublished doctoral dissertation, The University of Michigan, Ann Arbor, 1970.

French, E. G. Development of a measure of complex motivation. In Atkinson, J. W. (ed.). Motives in Fantasy, Action, and Society. Princeton: Van Nostrand, 1958.

French, E. G., and Lesser, G. S. Some characteristics of the achievement motive in women. Journal of Abnormal and Social Psychology, 68, 1964, 119-128.

Gurin, G., and Gurin, P. Expectancy theory in the study of poverty. Journal of Social Issues, 26, 1970, 83-104.

Gurin, P., Gurin, G., Lao, R., and Beattie, M. Internal-external control in the motivational dynamics of Negro youth. Journal of Social Issues, 25, 1969, 29-54.

Hagen, J. W. The effect of distraction on selective attention. Child Development, 38, 1967, 685-694.

Hagen, J., and Huntsman, N. Selective Attention in Mental Retardates. Paper presented at Society for Research in Child Development meeting. Los Angeles, 1969.

Hall, J. C. Correlation of a modified form of Raven's progressive matrices (1938) with the Wechsler Adult Intelligence Scale. Journal of Consulting Psychology, 21, 1957, 23-26.

Heckhausen, H. The Anatomy of Achievement Motivation. New York: Academic Press, 1967.

Horner, M. S. Sex Differences in Achievement Motivation and Performance in Competitive and Non-Competitive Situations. Unpublished doctoral dissertation, The University of Michigan, Ann Arbor, 1968.

Katz, I. Some motivational determinants of racial differences in intellectual achievement. International Journal of Psychology, 2, 1967, 1-12.

Kluckhohn, F. R. American women and American values. In Bryson, L. (ed.). Facing the Future's Risks. New York: Harper and Brothers, 1953, 175-199.

Lesser, G. S., Krawitz, R., and Packard, R. Experimental arousal of achievement motivation in adolscent girls. Journal of Abnormal and Social Psychology, 66, 1963, 59-66.

REFERENCES (Continued)

Lorge, I., and Thorndike, R. A. The Lorge-Thorndike Intelligence Tests. New York: Houghton Mifflin, 1950.

McClelland, D. C., Atkinson, J. W., Clark, R. A., and Lowell, E. L. The Achievement Motive. New York: Appleton-Century-Crofts, 1953.

Mannes, M. The problems of creative women. In Farber, S. M., and Wilson, R. H. L. (eds.). The Potential of Women. New York: McGraw-Hill, 1963.

Mead, M. Sex and Temperament in Three Primitive Societies. New York: William Morrow and Company, 1935.

Morgan, J. N., and Smith, J. D. A Panel Study of Income Dynamics: Study Design, Procedures, and Forms. Ann Arbor: Survey Research Center, Institute for Social Research, The University of Michigan, 1970.

Parsons, T. The social structure of the family. In Parsons, T. (ed.). The Family: Its Functions and Destiny. New York: Harper and Brothers, 1949.

Pierce, J. V., and Bowman, P. H. Motivational patterns of superior high school students. Cooperative Research Monographs, 1960, No. 2, 33-66. (USDHEW Publ. No. OE-35016). Cited in Lesser, G. S., Krawitz, R. and Packard, R., Experimental arousal of achievement motivation in adolescent girls. Journal of Abnormal and Social Psychology, 66, 1963, 59-66.

Raynor, J. O. Future orientation and motivation of immediate activity. Psychological Review, 76, 1969, 606-610.

Rogers, E. M., and Neill, R. E. Achievement Motivation Among Columbian Peasants. E. Lansing, Michigan: Department of Communication, Michigan State University, 1966.

Rotter, J. B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 80, 1966, No. 1 (whole No. 609), 1-28.

Sommerfeld, D. A. Job Training Programs in Detroit: A Comparative Study. Unpublished doctoral dissertation, The University of Michigan, Ann Arbor, 1969.

Veroff, J. A scoring manual for the power motive. In Atkinson, J. W. (ed.). Motives in Fantasy, Action, and Society. Princeton: Van Nostrand, 1958.

REFERENCES (Continued)

Veroff, J. Theoretical background for studying the origins of human motivational dispositions. Merrill-Palmer Quarterly, 11, 1965, 3-18.

Veroff, J., Atkinson, J. W., Feld, S. C., and Gurin, G. The use of thematic apperception to assess motivation in a nationwide interview study. Psychological Monographs, 1960, 94, No. 12 (whole no. 499).

Veroff, J. Social comparison and the development of achievement motivation. In Smith, C. D. (ed.), Achievement-Related Motives in Children. New York: Russell Sage, 1960.

Veroff, J., and Feld, S. C. Marriage and Work in America. New York: Van Nostrand-Reinhold, 1970.

Veroff, J., Feld, S. C., and Gurin, G. Dimensions of subjective adjustment. Journal of Abnormal and Social Psychology, 1962, 64, 192-205.

Veroff, J., and Peele, S. Initial effects of desegregation on the achievement motivation of Negro elementary school children. Journal of Social Issues, 1969, 25, 71-91.

Veroff, J., Wilcox, S., and Atkinson, J. W. The achievement motive in high school and college age women. Journal of Abnormal and Social Psychology, 1953, 48, 109-119.

Weiner, B. Effects of Unsatisfied Achievement-Related Motivation on Persistence and Subsequent Performance. Unpublished doctoral dissertation, The University of Michigan, Ann Arbor, 1963.